

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 21:00:56 ; Search time 50 Seconds  
(without alignments)  
885.694 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRWRTIYVGNLPGDIRK.....RSKSRSRSRNSPVPVWISG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1377.5	94.9	268	21 AAB11416	A. thaliana SRP30
2	1359.5	93.7	264	21 AAG47934	Arabidopsis thalia
3	1359.5	93.7	270	21 AAG23883	Arabidopsis thalia
4	1343	92.6	256	21 AAB11414	A. thaliana SRP30
5	1343	92.6	256	21 AAG47918	Arabidopsis thalia
6	1159	79.9	237	21 AAG30990	Arabidopsis thalia
7	865	59.6	303	21 AAG09881	Arabidopsis thalia
8	863	59.5	307	21 AAG19476	Arabidopsis thalia
9	836	57.6	303	21 AAB11417	A. thaliana SRP34/

10	805	55.5	296	21 AAG15601	Arabidopsis thalia
11	663.5	45.7	136	21 AAG47935	Arabidopsis thalia
12	663.5	45.7	142	21 AAG23884	Arabidopsis thalia
13	647	44.6	128	21 AAG47919	Arabidopsis thalia
14	646.5	44.6	294	21 AAG30418	Arabidopsis thalia
15	582.5	40.1	248	21 AAB11418	Human SF2/ASF prot
16	582.5	40.1	248	21 AAB90781	Human shear stress
17	547	37.7	255	22 ABB71532	Drosophila melanog
18	517	35.6	253	21 AAB57023	Human prostate can
19	517	35.6	254	23 ABA41316	Human ovarian anti
20	472	32.5	124	21 AAG12283	Zea mays protein f
21	472	32.5	159	21 AAG12282	Zea mays protein f
22	471.5	32.5	125	21 AAG35341	Zea mays protein f
23	471.5	32.5	154	21 AAG35340	Zea mays protein f
24	457.5	31.5	99	21 AAG09568	Arabidopsis thalia
25	453	31.2	227	22 AAU31465	Novel human secret
26	436	30.0	96	21 AAG47944	Arabidopsis thalia
27	431	29.7	329	22 ABB72060	Drosophila melanog
28	411.5	28.4	583	22 AAU30814	Novel human secret
29	391	25.9	83	21 AAG44000	Zea mays protein f
30	374.5	25.8	156	21 AAG27677	Arabidopsis thalia
31	363	25.0	165	21 AAG09882	Arabidopsis thalia
32	363	25.0	165	21 AAG19477	Arabidopsis thalia
33	356.5	24.6	155	21 AAG15602	Arabidopsis thalia
34	352	24.3	379	22 AAU30582	Novel human secret
35	339	23.4	153	22 AAU31463	Novel human secret
36	309	21.3	278	22 AAG00644	Novel human diagno
37	306	21.1	297	22 AAU30583	Novel human secret
38	277	19.1	306	21 AAB54149	Human pancreatic c
39	275.5	19.0	94	23 ABB31199	Human ORF172 prote
40	269.5	18.6	96	23 ABB00073	Human ORFX protein
41	268.5	18.5	117	21 AAG27678	Arabidopsis thalia
42	258.5	17.8	284	21 AAG05948	Arabidopsis thalia
43	258.5	17.8	313	21 AAG05947	Arabidopsis thalia
44	256	17.6	290	21 AAG05372	Arabidopsis thalia
45	254.5	17.5	283	22 AAM41305	Human polypeptide

ALIGNMENTS

RESULT 1  
AAB11416

ID AAB11416 standard; Protein; 268 AA.

XX AAB11416;

XX 23-FEB-2001 (first entry)

XX A. thaliana SRP30 protein.

XX SR protein; splice-factor activity; plant; developmental behavior;  
KW flowering; crop plant; cereal; bean; rice; fruit.

XX Arabidopsis thaliana.

PN WO200065059-A1.

XX 02-NOV-2000.

XX 20-APR-2000; 2000WO-AT00100.

XX 23-APR-1999; 99AT-0000727.

XX (OSTP ) OESTERR FORSCH SEIBERSDORF.

PI Barta A, Lopato S, Kalyana M, Dorner S;

XX WPI; 2000-687349/67.

PT Novel proteins with splice-factor activity in plants, useful e.g. for  
altering flowering time or development, and the nucleic acid that  
encodes it -

```
XX PS Disclosure; Fig 2; 67pp; German.
XX CC This invention describes a novel protein (I) with splice-factor activity
XX CC in plants (II) modifies the choice of splice sites in many plant
XX CC pre-mRNAs. (I) (also the nucleic acid that encodes them and related
XX CC vectors or expression systems) are used: (i) to alter splice patterns in
XX CC plants, or their parts; (ii) to alter developmental behavior of plants;
XX CC and/or (iii) to delay flowering, particularly by at least 25% relative
XX CC to the wild type, especially in crop plants such as cereals, beans, rice
XX CC and fruit.
XX SQ Sequence 268 AA;
    Query Match          94.9%; Score 1377.5; DB 21; Length 268;
    Best Local Similarity 96.1%; Pred. No. 8.5e-128;
    Matches 268; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
Qy   1  MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFEDPDADD 60
Db   1  MSSRWNTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFEDPDADD 60
Qy   61  AIYGRDGYDFDGCRLRVEIAHGRRFSPVDRYSSSYASRAPSRRSDYRVLVTGLPPSA 120
Db   61  AIYGRDGYDFDGCRLRVEIAHGRRFSPVDRYSSSYASRAPSRRSDYRVLVTGLPPSA 120
Qy   121  SWQDKDHRKAGDVCSEVPPDRKMGSGVVDYSDMKYAIKLDATFERNAPSSAYI 180
Db   121  SWQDKDHRKAGDVCSEVPPDRKMGSGVVDYSDMKYAIKLDATFERNAPSSAYI 180
Qy   181  RVRYEGRSVSRSPDDSKSVSRSGRSPGSCSKSRSVSPARSISPRSRPLSRGRSLY 240
Db   181  RVRYEGRSVSRSPDDSKSVSRSGRSPGSCSKSRSVSPARSISPRSRPLSRGRSLY 240
Qy   241  SSVSRSGSLRAGDWIQSRSKSRSRGRSRSNPSVSPVISG 279
Db   241  SSVSR-----SQSRSKSRSRGRSRSNPSVSPVISG 268

RESULT 2
AAG47934
ID  AAG47934 standard; Protein; 264 AA.
AC  AAG47934;
XX XX
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60472.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX XX
XX FN EP1033405-A2.
XX PD 06-SEP-2000.
XX XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127463.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-APR-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.
XX PR 18-JUN-1999; 99US-0139463.
XX PR 18-JUN-1999; 99US-0139750.
XX PR 18-JUN-1999; 99US-0139763.
XX PR 21-JUN-1999; 99US-0139817.
XX PR 22-JUN-1999; 99US-0139899.
XX PR 23-JUN-1999; 99US-0140353.
XX PR 23-JUN-1999; 99US-0140354.
XX PR 24-JUN-1999; 99US-0140695.
XX PR 28-JUN-1999; 99US-0140823.
XX PR 29-JUN-1999; 99US-0140991.
XX PR 30-JUN-1999; 99US-0141287.
XX PR 01-JUL-1999; 99US-0141842.
XX PR 01-JUL-1999; 99US-0142154.
XX PR 02-JUL-1999; 99US-0142055.
XX PR 06-JUL-1999; 99US-0142390.
XX PR 08-JUL-1999; 99US-0142803.
XX PR 09-JUL-1999; 99US-0142920.
XX PR 12-JUL-1999; 99US-0142977.
XX PR 13-JUL-1999; 99US-0143542.
XX PR 14-JUL-1999; 99US-0143624.
XX PR 15-JUL-1999; 99US-0144005.
XX PR 16-JUL-1999; 99US-0144085.
XX PR 16-JUL-1999; 99US-0144086.
XX PR 19-JUL-1999; 99US-0144325.
XX PR 19-JUL-1999; 99US-0144331.
XX PR 19-JUL-1999; 99US-0144332.
XX PR 19-JUL-1999; 99US-0144333.
XX PR 19-JUL-1999; 99US-0144334.
XX PR 19-JUL-1999; 99US-0144335.
XX PR 20-JUL-1999; 99US-0144352.
```

PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	23-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	25-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
PR	07-SEP-1999;	99US-0152363.			
PR	10-SEP-1999;	99US-0153070.			
PR	13-SEP-1999;	99US-0153758.			
PR	15-SEP-1999;	99US-0154018.			
PR	16-SEP-1999;	99US-0154039.			
PR	22-SEP-1999;	99US-0154779.			
PR	22-SEP-1999;	99US-0155139.			
PR	23-SEP-1999;	99US-0155486.			
PR	24-SEP-1999;	99US-0155659.			
PR	28-SEP-1999;	99US-0156458.			
PR	29-SEP-1999;	99US-0156596.			
PR	04-OCT-1999;	99US-0157117.			
PR	05-OCT-1999;	99US-0157753.			
PR	06-OCT-1999;	99US-0157865.			
PR	07-OCT-1999;	99US-0158029.			
PR	08-OCT-1999;	99US-0158232.			
PR	12-OCT-1999;	99US-0158369.			
PR	13-OCT-1999;	99US-0159293.			
PR	13-OCT-1999;	99US-0159294.			
PR	13-OCT-1999;	99US-0159295.			
PR	14-OCT-1999;	99US-0159329.			
PR	14-OCT-1999;	99US-0159330.			
PR	14-OCT-1999;	99US-0159331.			
Query Match 93.7%; Score 1359.5; DB 21; Length 264;					
Best Local Similarity 96.0%; Pred. No. 5e-126; Mismatches 0; Indels 11; Gaps 1;					
Matches 264; Conservative					
QY	1	MSRWNRRTIYVGNLPGDIRKCEVDLFYKYGPVIVDILKIPRPPGYAFVEFEDPRDADD	60		
Db	1	MSRWNRRTIYVGNLPGDIRKCEVDLFYKYGPVIVDILKIPRPPGYAFVEFEDPRDADD	60		
QY	61	AIYGRDGYDFDQCLRLVEIAHGRRFSPSVDRYSSYSASRAPSRSDYELVTGLPPSA	120		
Db	61	AIYGRDGYDFDQCLRLVEIAHGRRFSPSVDRYSSYSASRAPSRSDYELVTGLPPSA	120		
QY	121	SWQDLKDHMRKAGDVCFSEVFPDRKMGVVDYSDYNDMKYAIRKLDATFENAFSSAYI	180		
Db	121	SWQDLKDHMRKAGDVCFSEVFPDRKMGVVDYSDYNDMKYAIRKLDATFENAFSSAYI	180		
QY	181	RVREYESRSVSRSPDDSKYSRSGPSCSYSSKRSVSPARSISPRSLRSRSY	240		
Db	181	RVREYESRSVSRSPDDSKYSRSGPSCSYSSKRSVSPARSISPRSLRSRSY	240		
QY	241	SSVSRSGSLLRAGDWISQSRKSRGRSRNSPVP	275		
Db	241	SSVSR-----SQSRKSRGRSRNSPVP	264		
RESULT 3					
AAG23883					
ID	AAG23883 standard; Protein; 270 AA.				
XX	AAG23883;				
AC					
XX	17-OCT-2000 (first entry)				
DT					
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 27353.				
DE	Arabidopsis thaliana				
XX	Protein identification; signal transduction pathway; metabolic pathway;				
KW	Hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX	Arabidopsis thaliana.				
OS					
XX	EP1033405-A2.				
PN					
XX	06-SEP-2000.				
XX	25-FEB-2000; 2000EP-0301439.				
PF					
XX	25-FEB-1999; 99US-0121825.				
PR	05-MAR-1999; 99US-0123180.				
PR	09-MAR-1999; 99US-0123548.				

PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134258.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139483.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142997.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.



```
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      93.7%; Score 1359.5; DB 21; Length 270;
Best Local Similarity 95.0%; Pred. No. 5.2e-126;
Matches 267; Conservative 0; Mismatches 1; Indels 13; Gaps 2;

QY 1 MSSRWNRITVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Db 1 MSSRWNRITVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
QY 61 AIYGRDGYDFDGCRLRVEIAHGRRPSPVDYSSYSASRAPSRSDYRVLTGLPPSA 120
Db 61 AIYGRDGYDFDGCRLRVEIAHGRRPSPVDYSSYSASRAPSRSDYRVLTGLPPSA 120
QY 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSNYDDMKYAIRKLDATFERNAFSSAYI 180
Db 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSNYDDMKYAIRKLDATFERNAFSSAYI 180
QY 181 RVREYESRSVSRSPDDSKYSRSPSRGPGSCSYSSKSRSPARSTSPSRPLSRSRSLY 240
Db 181 RVREYESRSVSRSPDDSKYSRSPSRGPGSCSYSSKSRSPARSTSPSRPLSRSRSPY 240
QY 241 SSVSRSGSLLRAGDWI 256
Db 241 SSVSRSGSLLRAGDWI 256
```

```
RESULT 4
AAB11414
ID AAB11414 standard; Protein; 256 AA.
XX
AC AAB11414;
XX
XX 23-FEB-2001 (first entry)
XX
XX A. thaliana SRp30 protein.
XX
XX SR protein; splice-factor activity; plant; developmental behavior;
XX flowering; crop plant; cereal; bean; rice; fruit.
XX
XX Arabidopsis thaliana.
XX
XX WO200065059-A1.
```

```
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-AT00100.
XX
XX 23-APR-1999; 99AT-0000727.
XX
XX (OSTP ) OESTERR FORSCH SEIBERSDORF.
XX
XX Barta A, Lopato S, Kalyna M, Dorner S;
XX WPI; 2000-687349/67.
XX N-PSDB; AAC81899.
XX
XX Novel proteins with splice-factor activity in plants, useful e.g. for
XX altering flowering time or development, and the nucleic acid that
XX encodes it -
XX
XX Claim 1; Fig 1A; 67pp; German.
XX
XX This invention describes a novel protein (I) with splice-factor activity
XX in plants (I) modifies the choice of splice sites in many plant
XX pre-mRNAs. (I) (also the nucleic acid that encodes them and related
XX vectors or expression systems) are used: (i) to alter splice patterns in
XX plants, or their parts; (ii) to alter developmental behavior of plants;
XX and/or (iii) to delay flowering, particularly by at least 25% relative
XX to the wild type, especially in crop plants such as cereals, beans, rice
XX and fruit.
XX
XX Sequence 256 AA;

Query Match      92.6%; Score 1343; DB 21; Length 256;
Best Local Similarity 100.0%; Pred. No. 2.1e-124;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRWNRITVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
Db 1 MSSRWNRITVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60
QY 61 AIYGRDGYDFDGCRLRVEIAHGRRPSPVDYSSYSASRAPSRSDYRVLTGLPPSA 120
Db 61 AIYGRDGYDFDGCRLRVEIAHGRRPSPVDYSSYSASRAPSRSDYRVLTGLPPSA 120
QY 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSNYDDMKYAIRKLDATFERNAFSSAYI 180
Db 121 SWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSNYDDMKYAIRKLDATFERNAFSSAYI 180
QY 181 RVREYESRSVSRSPDDSKYSRSPSRGPGSCSYSSKSRSPARSTSPSRPLSRSRSLY 240
Db 181 RVREYESRSVSRSPDDSKYSRSPSRGPGSCSYSSKSRSPARSTSPSRPLSRSRSLY 240
QY 241 SSVSRSGSLLRAGDWI 256
Db 241 SSVSRSGSLLRAGDWI 256

RESULT 5
AAG47918
ID AAG47918 standard; Protein; 256 AA.
XX
XX AAG47918;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60450.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
```

XX	06-SEP-2000.		PR	02-JUL-1999;	99US-0142055.
PD			PR	06-JUL-1999;	99US-0142390.
PF	25-FEB-2000;	2000EP-0301439.	PR	08-JUL-1999;	99US-0142803.
XX			PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-1999;	99US-0121825.	PR	12-JUL-1999;	99US-0142977.
PR	05-MAR-1999;	99US-0123180.	PR	13-JUL-1999;	99US-0143542.
PR	09-MAR-1999;	99US-0123548.	PR	14-JUL-1999;	99US-0143621.
PR	23-MAR-1999;	99US-0125788.	PR	15-JUL-1999;	99US-0144005.
PR	25-MAR-1999;	99US-0126264.	PR	16-JUL-1999;	99US-0144085.
PR	29-MAR-1999;	99US-0126785.	PR	16-JUL-1999;	99US-0144086.
PR	01-APR-1999;	99US-0126785.	PR	19-JUL-1999;	99US-0144325.
PR	06-APR-1999;	99US-0127482.	PR	19-JUL-1999;	99US-0144331.
PR	08-APR-1999;	99US-0128234.	PR	19-JUL-1999;	99US-0144332.
PR	16-APR-1999;	99US-0128714.	PR	19-JUL-1999;	99US-0144333.
PR	19-APR-1999;	99US-0129845.	PR	19-JUL-1999;	99US-0144334.
PR	21-APR-1999;	99US-0130077.	PR	19-JUL-1999;	99US-0144335.
PR	23-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144684.
PR	30-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135623.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137282.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139482.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139455.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139456.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139457.	PR	18-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139458.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139461.	PR	20-AUG-1999;	99US-0149829.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139463.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139750.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139763.	PR	25-AUG-1999;	99US-0150566.
PR	21-JUN-1999;	99US-0139817.	PR	26-AUG-1999;	99US-0150884.
PR	22-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0151065.
PR	23-JUN-1999;	99US-0140353.	PR	27-AUG-1999;	99US-0151066.
PR	23-JUN-1999;	99US-0140354.	PR	27-AUG-1999;	99US-0151080.
PR	24-JUN-1999;	99US-0140695.	PR	30-AUG-1999;	99US-0151303.
PR	28-JUN-1999;	99US-0140823.	PR	31-AUG-1999;	99US-0151438.
PR	29-JUN-1999;	99US-0140991.	PR	01-SEP-1999;	99US-0151930.
PR	30-JUN-1999;	99US-0141287.	PR	07-SEP-1999;	99US-0152363.
PR	01-JUL-1999;	99US-0141842.	PR	10-SEP-1999;	99US-0153070.
PR	01-JUL-1999;	99US-0142154.	PR	15-SEP-1999;	99US-0153758.
			PR	16-SEP-1999;	99US-0154018.
			PR	20-SEP-1999;	99US-0154039.
			PR		99US-0154779.



PR 22-JUN-1999;	99US-0139899;	PR 30-AUG-1999;	99US-0151303;
PR 23-JUN-1999;	99US-0140353;	PR 31-AUG-1999;	99US-0151438;
PR 23-JUN-1999;	99US-0140354;	PR 01-SEP-1999;	99US-0151930;
PR 24-JUN-1999;	99US-0140695;	PR 07-SEP-1999;	99US-0152363;
PR 28-JUN-1999;	99US-0140823;	PR 10-SEP-1999;	99US-0153070;
PR 29-JUN-1999;	99US-0140991;	PR 13-SEP-1999;	99US-0153758;
PR 30-JUN-1999;	99US-0141287;	PR 15-SEP-1999;	99US-0154018;
PR 01-JUL-1999;	99US-0141842;	PR 20-SEP-1999;	99US-0154039;
PR 01-JUL-1999;	99US-0142154;	PR 20-SEP-1999;	99US-0154779;
PR 02-JUL-1999;	99US-0142055;	PR 22-SEP-1999;	99US-0155139;
PR 06-JUL-1999;	99US-0142390;	PR 23-SEP-1999;	99US-0155486;
PR 08-JUL-1999;	99US-0142803;	PR 24-SEP-1999;	99US-0156659;
PR 09-JUL-1999;	99US-0142920;	PR 28-SEP-1999;	99US-0156458;
PR 12-JUL-1999;	99US-0142977;	PR 29-SEP-1999;	99US-0156596;
PR 13-JUL-1999;	99US-0143542;	PR 04-OCT-1999;	99US-0157117;
PR 14-JUL-1999;	99US-0143624;	PR 05-OCT-1999;	99US-0157753;
PR 15-JUL-1999;	99US-0144005;	PR 06-OCT-1999;	99US-0157865;
PR 16-JUL-1999;	99US-0144085;	PR 07-OCT-1999;	99US-0158029;
PR 16-JUL-1999;	99US-0144086;	PR 08-OCT-1999;	99US-0158232;
PR 19-JUL-1999;	99US-0144325;	PR 12-OCT-1999;	99US-0158369;
PR 19-JUL-1999;	99US-0144331;	PR 13-OCT-1999;	99US-0159293;
PR 19-JUL-1999;	99US-0144332;	PR 13-OCT-1999;	99US-0159294;
PR 19-JUL-1999;	99US-0144333;	PR 13-OCT-1999;	99US-0159295;
PR 19-JUL-1999;	99US-0144334;	PR 14-OCT-1999;	99US-0159295;
PR 19-JUL-1999;	99US-0144335;	PR 14-OCT-1999;	99US-0159330;
PR 20-JUL-1999;	99US-0144352;	PR 14-OCT-1999;	99US-0159637;
PR 20-JUL-1999;	99US-0144632;	PR 14-OCT-1999;	99US-0159638;
PR 20-JUL-1999;	99US-0144884;	PR 18-OCT-1999;	99US-0159584;
PR 21-JUL-1999;	99US-0144814;	PR 21-OCT-1999;	99US-0160741;
PR 21-JUL-1999;	99US-0145086;	PR 21-OCT-1999;	99US-0160767;
PR 21-JUL-1999;	99US-0145088;	PR 21-OCT-1999;	99US-0160768;
PR 22-JUL-1999;	99US-0145085;	PR 21-OCT-1999;	99US-0160770;
PR 22-JUL-1999;	99US-0145087;	PR 21-OCT-1999;	99US-0160814;
PR 22-JUL-1999;	99US-0145089;	PR 21-OCT-1999;	99US-0160815;
PR 22-JUL-1999;	99US-0145192;	PR 22-OCT-1999;	99US-0160980;
PR 23-JUL-1999;	99US-0145145;	PR 22-OCT-1999;	99US-0160981;
PR 23-JUL-1999;	99US-0145218;	PR 22-OCT-1999;	99US-0160989;
PR 23-JUL-1999;	99US-0145224;	PR 25-OCT-1999;	99US-0161404;
PR 26-JUL-1999;	99US-0145276;	PR 25-OCT-1999;	99US-0161405;
PR 27-JUL-1999;	99US-0145913;	PR 25-OCT-1999;	99US-0161406;
PR 27-JUL-1999;	99US-0145918;	PR 26-OCT-1999;	99US-0161359;
PR 27-JUL-1999;	99US-0145919;	PR 26-OCT-1999;	99US-0161360;
PR 28-JUL-1999;	99US-0145951;	PR 26-OCT-1999;	99US-0161361;
PR 02-AUG-1999;	99US-0146386;	PR 28-OCT-1999;	99US-0161920;
PR 02-AUG-1999;	99US-0146388;	PR 28-OCT-1999;	99US-0161992;
PR 02-AUG-1999;	99US-0146399;	PR 28-OCT-1999;	99US-0161993;
PR 03-AUG-1999;	99US-0147038;	PR 29-OCT-1999;	99US-0162142;
PR 04-AUG-1999;	99US-0147204;		
PR 04-AUG-1999;	99US-0147302;		
PR 05-AUG-1999;	99US-0147192;		
PR 05-AUG-1999;	99US-0147260;		
PR 06-AUG-1999;	99US-0147303;		
PR 06-AUG-1999;	99US-0147416;		
PR 09-AUG-1999;	99US-0147493;		
PR 09-AUG-1999;	99US-0147935;		
PR 10-AUG-1999;	99US-0148171;		
PR 11-AUG-1999;	99US-0148319;		
PR 12-AUG-1999;	99US-0148341;		
PR 13-AUG-1999;	99US-0148565;		
PR 13-AUG-1999;	99US-0148684;		
PR 16-AUG-1999;	99US-0149368;		
PR 17-AUG-1999;	99US-0149175;		
PR 18-AUG-1999;	99US-0149426;		
PR 20-AUG-1999;	99US-0149722;		
PR 20-AUG-1999;	99US-0149723;		
PR 20-AUG-1999;	99US-0149929;		
PR 23-AUG-1999;	99US-0149902;		
PR 23-AUG-1999;	99US-0149930;		
PR 26-AUG-1999;	99US-0150566;		
PR 26-AUG-1999;	99US-0150884;		
PR 27-AUG-1999;	99US-0151065;		
PR 27-AUG-1999;	99US-0151066;		
PR 27-AUG-1999;	99US-0151080;		

Query Match

Best Local Similarity

Matches 218; Conservative

79.9%; Score 1159; DB 21; Length 237;

100.0%; Pred. No. 3e-106;

0; Mismatches

0; Indels

0; Gaps

0;

QY	1	MSSRWNRITTYGNLPGDIRKCEVEDLFYKYGPIVDIDKIPRPPGVAFFEDPRDADD	60
Db	1	MSSRWNRITTYGNLPGDIRKCEVEDLFYKYGPIVDIDKIPRPPGVAFFEDPRDADD	60
QY	61	AIYGRDGDYDFGCRRLRVBIANGRRFSPSDRYSSYSASAPRRSDYRVLTGLPESA	120
Db	61	AIYGRDGDYDFGCRRLRVBIANGRRFSPSDRYSSYSASAPRRSDYRVLTGLPESA	120
QY	121	SWQDLKDHMRKAGDVCVFSEVPDRKMGVVDYNDYDMKYAIRKLDATFRNAPSSAYI	180
Db	121	SWQDLKDHMRKAGDVCVFSEVPDRKMGVVDYNDYDMKYAIRKLDATFRNAPSSAYI	180
QY	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKSR	218
Db	181	RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKSR	218

RESULT 7  
AAG09881  
ID AAG09881 standard; Protein; 303 AA.

XX	AAG09881;	PR	18-JUN-1999;	99US-0139462.
AC		PR	18-JUN-1999;	99US-0139463.
XX		PR	18-JUN-1999;	99US-0139750.
DT	17-OCT-2000 (first entry)	PR	18-JUN-1999;	99US-0139763.
XX		PR	21-JUN-1999;	99US-0139817.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 7985.	PR	22-JUN-1999;	99US-0139899.
XX		PR	23-JUN-1999;	99US-0140353.
KW	Protein identification; signal transduction pathway; metabolic pathway;	PR	23-JUN-1999;	99US-0140354.
KW	hybridisation assay; Genetic mapping; gene expression control; promoter;	PR	24-JUN-1999;	99US-0140695.
KW	termination sequence.	PR	28-JUN-1999;	99US-0140823.
XX		PR	29-JUN-1999;	99US-0140991.
OS	Arabidopsis thaliana.	PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144684.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148585.
PR	17-JUN-1999;	PR	16-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149168.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.



PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144086.
PR	16-JUL-1999;	99US-0144085.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145192.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148565.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	23-AUG-1999;	99US-014930.
PR	23-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	13-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	28-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161820.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 59.5%; Score 863; DB 21; Length 307;		
Best Local Similarity 62.2%; Fred. No. 8.3e-77;		
Matches 184; Conservative 33; Mismatches 45; Indels 34; Gaps 7;		
Qy	1	MSSRWRTIVVGNLPGDIRKCEVEDLFYKYGPVTDLKIIPRPPGYAFVEFEDPRDAD 60
Db	1	MSSRSRTIVVGNLPGDIRREVEDFSKIGPVQDILKVPDPGGYAFVEFDDARDAED 60
Qy	61	AIYGRDGYDFDGCRLRVEIAHGGRFSPSVDRYSSYS-----ASRAP 103

Db 61 AIHGRDGYDFDCHRLRVELAHGRR---SSDDTRGSENGGRRGGRRGGRRGGRRGGRRGG 117

Qy 104 SRRSDYRLVTGLPPSSASQDLKHMEKAGDVCFSEVFPPDRKMGSGVVDYNYDDMKYAI 163

Db 118 SRRSEFRVLVTGLPSSASQDLKHMEKGGDVCFSQVYRDARTGTGVDVTCYEDMKYAL 177

Qy 164 KLDATFRNFAFSAIYRVREYESSRSVSRSPDDSKSY-RSRSRGRGSPCSYSSKSRSPV 222

Db 178 KCLDTEFRNFAFNGYVREYDSRKDSRSPRGESYSKSRSRGRSV---SRSRSRSR 234

Qy 223 ARSTSPSRPLRSRLSYSSVSRGSLLRAGDWISQSRKS--RSRGRNSPVPV 276

Db 235 SRRSPKAK--SSRRSPAKSTSRSPGPR-----SKSRSPRRSRSRSPPLSPV 282

RESULT 9

RAB11417

ID AAB11417 standard; Protein; 303 AA.

XX

AC AAB11417;

XX

DT 23-FEB-2001 (first entry)

XX

DE A. thaliana SRP34/SRI protein.

XX

KW SR protein; splice-factor activity; plant; developmental behavior;

KW flowering; crop plant; cereal; bean; rice; fruit.

XX

OS Arabidopsis thaliana.

XX

PN WO200605059-A1.

XX

PD 02-NOV-2000.

XX

PF 20-APR-2000; 2000WO-AT00100.

XX

PR 23-APR-1999; 59AT-0000727.

PA (OSTP ) OESTERR FORSCH SETBERSDORF.

PI Barta A, Lopato S, Kalyna M, Dornier S;

XX

DR WPI; 2000-687349/67.

XX

PT Novel proteins with splice-factor activity in plants, useful e.g. for

PT altering flowering time or development, and the nucleic acid that

PT encodes it -

XX

PS Disclosure; Fig 2; 67pp; German.

XX

CC This invention describes a novel protein (I) with splice-factor activity

CC in plants (I) modifies the choice of splice sites in many plant

CC pre-mRNAs. (I) (also the nucleic acid that encodes them and related

CC vectors or expression systems) are used: (i) to alter splice patterns in

CC plants, or their parts; (ii) to alter developmental behavior of plants;

CC and/or (iii) to delay flowering, particularly by at least 25% relative

CC to the wild type, especially in crop plants such as cereals, beans, rice

CC and fruit.

XX

SQ Sequence 303 AA;

Query Match 57.6%; Score 836; DB 21; Length 303;

Best Local Similarity 62.0%; Pred. No. 3.8e-74;

Matches 181; Conservative 32; Mismatches 49; Indels 30; Gaps 7;

Qy 1 MSSRNRTIYGNLPGDKKEVEDLFVYKGPVIDIDIKIPRPPGYAFVEFEDPRADD 60

Db 1 MSSRSRTIYGNLPGDKIREVEDLFKYPVQIDILKVPFRPPGYAFVEFEDDARDED 60

Qy 61 AIYGRDGYDFDCHRLRVELAHGRR---SSDDTRGSENGGRRGGRRGGRRGGRRGG 107

Db 61 AIHGRDGYDFDCHRLRVELAHGRR---SSDDTRGSENGGRRGGRRGGRRGGRRGG 117

Qy 108 DYRLVLTGLPPSSASQDLKHMEKAGDVCFSEVFPPDRKMGSGVVDYNYDDMKYAIKLD 167

Db 118 EERVLVTWLASSASQDLKHIAKGGDVCFSQVYRDARTGTGVDVTCYEDMKYALKLD 177

Qy 168 ATEFRNFAFSAIYRVREYESSRSVSRSPDDSKSY-RSRSRGRGSPCSYSSKSRSPV 226

Db 178 DTEFRNFAFNGYVREYDSRKDSRSPRGESYSKSRSRGRSV---SRSRSRSRSR 234

Qy 227 SPRSRPLRSRLSYSSVSRGSLLRAGDWISQSRKS--RSRGRNSPVPV 276

Db 235 SPKAK--SSRRSPAKSTSRSPGPR-----SKSRSPRRSRSRSPPLSPV 278

RESULT 10

AAG15601

ID AAG15601 standard; Protein; 296 AA.

XX

AC AAG15601;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 15917.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; Genetic mapping; gene expression control; promoter;

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0132866.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.



PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147152.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156438.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	22-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 55.5%; Score 805; DB 21; Length 296;		
Best Local Similarity 62.5%; Pred. No. 4.3e-71;		
Matches 183; Conservative 27; Mismatches 53; Indels 30; Gaps 10;		

Qy	1	MSRNRWNTIVGNLPGDIRKCEVEDLKYKIGPIVDIDLKIPRPPGYATVEBEDPRDADD	60
Db	1	MSGRFERSIYVGNLPGDIRHEHEDIFYKYGRIVDIELKVPPEPPCYCFVEFHSRDAED	60
Qy	61	AIYGRDGYDDGRLRVEIAHGGRFSPVDRY-----SSSY-----SASRAPSR	106
Db	61	AIKRGDYNLDGRLRVELAHGGRGQSSS-DRRGYGGGGGYGGGGGGGSGARFGVSRH	119
Qy	107	SDYRVLTGLPPSASWQDLKDHMRKAGDVCFSVFPDRKGMGVVDYSNYDDMKYAIRKL	166
Db	120	SEFVIVRGLPSSASWQDLKDHMRKAGDVCFAEVTDSGTGYGVVDYTNVDDMKYAIRKL	179
Qy	167	DATEFRNAFSAVIRVREYE-SRSVSRSPDDSKYSRSGRSPCSYS-SKSRSVSPAR	224
Db	180	DDTEFRNPWARGIRVKYKESRSRSPSRSS-RSRSGRGRSHSRSLSRSPK	238
Qy	225	SI--SPR--SRPLSRSLYSVSRSGSLLRAGDWISQSRKSRSRSPNSP	272
Db	239	DLKSPRSLSRSGISRS--PSPDKKPPRA-----MSRKSRSRSP	284
RESULT 11			
AAG47935			
XX	ID	AAG47935 standard; Protein; 136 AA.	
XX	AC	AAG47935;	
XX	XX		
DT	18-OCT-2000	(first entry)	
XX	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 60473.	
DE	DE	Protein identification; signal transduction pathway; metabolic pathway;	
KW	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	KW	termination sequence.	
XX	XX	Arabidopsis thaliana.	
XX	XX	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	XX	25-FEB-2000; 2000EF-0301439.	
XX	XX		
PR	25-FEB-1999;	99US-0121825.	
PR	03-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127452.	
PR	06-APR-1999;	99US-0128234.	
PR	16-APR-1999;	99US-0128714.	
PR	19-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	

PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	10-JUN-1999;	99US-0138847.	
PR	14-JUN-1999;	99US-0139119.	
PR	16-JUN-1999;	99US-0139452.	
PR	16-JUN-1999;	99US-0139453.	
PR	17-JUN-1999;	99US-0139492.	
PR	18-JUN-1999;	99US-0139454.	
PR	18-JUN-1999;	99US-0139455.	
PR	18-JUN-1999;	99US-0139456.	
PR	18-JUN-1999;	99US-0139457.	
PR	18-JUN-1999;	99US-0139458.	
PR	18-JUN-1999;	99US-0139459.	
PR	18-JUN-1999;	99US-0139460.	
PR	18-JUN-1999;	99US-0139461.	
PR	18-JUN-1999;	99US-0139462.	
PR	18-JUN-1999;	99US-0139463.	
PR	18-JUN-1999;	99US-0139750.	
PR	18-JUN-1999;	99US-0139763.	
PR	21-JUN-1999;	99US-0139817.	
PR	22-JUN-1999;	99US-0139899.	
PR	23-JUN-1999;	99US-0140353.	
PR	23-JUN-1999;	99US-0140354.	
PR	24-JUN-1999;	99US-0140695.	
PR	28-JUN-1999;	99US-0140823.	
PR	29-JUN-1999;	99US-0140931.	
PR	30-JUN-1999;	99US-0141287.	
PR	01-JUL-1999;	99US-0141842.	
PR	01-JUL-1999;	99US-0142154.	
PR	02-JUL-1999;	99US-0142055.	
PR	06-JUL-1999;	99US-0142390.	
PR	08-JUL-1999;	99US-0142803.	
PR	09-JUL-1999;	99US-0142920.	
PR	12-JUL-1999;	99US-0142977.	
PR	13-JUL-1999;	99US-0143542.	
PR	14-JUL-1999;	99US-0143624.	
PR	15-JUL-1999;	99US-0144005.	
PR	16-JUL-1999;	99US-0144085.	
PR	16-JUL-1999;	99US-0144086.	
PR	19-JUL-1999;	99US-0144325.	
PR	19-JUL-1999;	99US-0144331.	
PR	19-JUL-1999;	99US-0144332.	
PR	19-JUL-1999;	99US-0144333.	
PR	19-JUL-1999;	99US-0144334.	
PR	19-JUL-1999;	99US-0144335.	
PR	20-JUL-1999;	99US-0144352.	
PR	20-JUL-1999;	99US-0144632.	
PR	20-JUL-1999;	99US-0144884.	
PR	21-JUL-1999;	99US-0144814.	
PR	21-JUL-1999;	99US-0145086.	
PR	21-JUL-1999;	99US-0145088.	
PR	22-JUL-1999;	99US-0145085.	
PR	22-JUL-1999;	99US-0145087.	
PR	22-JUL-1999;	99US-0145089.	
PR	22-JUL-1999;	99US-0145192.	
PR	23-JUL-1999;	99US-0145145.	
PR	23-JUL-1999;	99US-0145218.	
PR	23-JUL-1999;	99US-0145224.	
PR	26-JUL-1999;	99US-0145276.	
PR	27-JUL-1999;	99US-0145913.	
PR	27-JUL-1999;	99US-0145918.	
PR	27-JUL-1999;	99US-0145919.	
PR	28-JUL-1999;	99US-0145951.	

```

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

Query Match 45.7%; Score 663.5; DB 21; Length 136;
Best Local Similarity 92.5%; Pred. No. 1.4e-57;
Matches 136; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

Qy 129 MRKAGDVCFSEVFPDRKGMGVVDYSGYSDMKYAIKRLDATEFRNAPSSAYIRVREYER 188
Db 1 MRKAGDVCFSEVFPDRKGMGVVDYSGYSDMKYAIKRLDATEFRNAPSSAYIRVREYER 60

Qy 189 SVSRSPDDSKSYRSRSGRSGPSCSYSSKRSVSPPARSISPRSRPLSRSLYSSVSRSGS 248
Db 61 SVSRSPDDSKSYRSRSGRSGPSCSYSSKRSVSPPARSISPRSRPLSRSLYSSVSRSGS 117

Qy 249 LLRAGDMWISQSRSKRSRSGRSGSPVSP 275
Db 118 -----SQSRSKRSRSGRSGSPVSP 136

RESULT 12
AAG23884
ID AAG23884 standard; Protein; 142 AA.
XX AC AAG23884;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 27354.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX EN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123348.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.

```

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135639.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140655.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.

```
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 45.7%; Score 663.5; DB 21; Length 142;
Best Local Similarity 90.8%; Pred. No. 1.5e-57;
Matches 139; Conservative 0; Mismatches 1; Indels 13; Gaps 2;

QY 129 MRKAGDVCSEVFPDRKMGVVDYNSYDDMKYAIKLDATFERNAFSAVIRVEYSR 188
DB 1 MRKAGDVCSEVFPDRKMGVVDYNSYDDMKYAIKLDATFERNAFSAVIRVEYSR 60

QY 189 SVSRSPDDSKYRSRSGRGPSCSYSSKRSRVSYPARSISPRSRPLSRGRSLYSSVSRSGS 248
DB 61 SVSRSPDDSKYRSRSGRGPSCSYSSKRSRVSYPARSISPRSRPLSRGRSLYSSVSR --- 117

QY 249 LLRAGDWTISQSRSKRS--RSRNSPVPSPVSG 279
DB 118 -----SQSRSKRSTXRSRNSPVPSPVSG 142

RESULT 13
AAG47919
ID AAG47919 standard; Protein; 128 AA.
AC AAG47919;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60451.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
```

27-JUL-1999; 99US-0145919.  
 28-JUL-1999; 99US-0145951.  
 28-AUG-1999; 99US-0146386.  
 02-AUG-1999; 99US-0146388.  
 02-AUG-1999; 99US-0146389.  
 03-AUG-1999; 99US-0147038.  
 04-AUG-1999; 99US-0147204.  
 04-AUG-1999; 99US-0147302.  
 05-AUG-1999; 99US-0147192.  
 05-AUG-1999; 99US-0147260.  
 06-AUG-1999; 99US-0147303.  
 06-AUG-1999; 99US-0147416.  
 09-AUG-1999; 99US-0147493.  
 09-AUG-1999; 99US-0147935.  
 10-AUG-1999; 99US-0148171.  
 11-AUG-1999; 99US-0148319.  
 12-AUG-1999; 99US-0148341.  
 13-AUG-1999; 99US-0148565.  
 13-AUG-1999; 99US-0148684.  
 16-AUG-1999; 99US-0149368.  
 17-AUG-1999; 99US-0149175.  
 18-AUG-1999; 99US-0149426.  
 20-AUG-1999; 99US-0149722.  
 20-AUG-1999; 99US-0149723.  
 20-AUG-1999; 99US-0149929.  
 23-AUG-1999; 99US-0149902.  
 23-AUG-1999; 99US-0149930.  
 25-AUG-1999; 99US-0150566.  
 26-AUG-1999; 99US-0150884.  
 27-AUG-1999; 99US-0151065.  
 27-AUG-1999; 99US-0151066.  
 27-AUG-1999; 99US-0151080.  
 30-AUG-1999; 99US-0151303.  
 31-AUG-1999; 99US-0151438.  
 01-SEP-1999; 99US-0151930.  
 07-SEP-1999; 99US-0152363.  
 10-SEP-1999; 99US-0153070.  
 13-SEP-1999; 99US-0153758.  
 15-SEP-1999; 99US-0154018.  
 16-SEP-1999; 99US-0154039.  
 20-SEP-1999; 99US-0154779.  
 22-SEP-1999; 99US-0155139.  
 23-SEP-1999; 99US-0155486.  
 24-SEP-1999; 99US-0155659.  
 28-SEP-1999; 99US-0156458.  
 29-SEP-1999; 99US-0156596.  
 04-OCT-1999; 99US-0157117.  
 05-OCT-1999; 99US-0157753.  
 06-OCT-1999; 99US-0157865.  
 07-OCT-1999; 99US-0158029.  
 08-OCT-1999; 99US-0158232.  
 12-OCT-1999; 99US-0158369.  
 13-OCT-1999; 99US-0159293.  
 13-OCT-1999; 99US-0159294.  
 13-OCT-1999; 99US-0159295.  
 14-OCT-1999; 99US-0159329.  
 14-OCT-1999; 99US-0159330.  
 14-OCT-1999; 99US-0159331.  
 14-OCT-1999; 99US-0159637.  
 14-OCT-1999; 99US-0159638.  
 18-OCT-1999; 99US-0159584.  
 21-OCT-1999; 99US-0160741.  
 21-OCT-1999; 99US-0160767.  
 21-OCT-1999; 99US-0160768.  
 21-OCT-1999; 99US-0160770.  
 21-OCT-1999; 99US-0160814.  
 21-OCT-1999; 99US-0160815.  
 22-OCT-1999; 99US-0160980.  
 22-OCT-1999; 99US-0160981.  
 22-OCT-1999; 99US-0160989.  
 25-OCT-1999; 99US-0161404.  
 25-OCT-1999; 99US-0161405.  
 25-OCT-1999; 99US-0161406.

26-OCT-1999; 99US-0161359.  
 26-OCT-1999; 99US-0161360.  
 26-OCT-1999; 99US-0161361.  
 28-OCT-1999; 99US-0161920.  
 28-OCT-1999; 99US-0161922.  
 28-OCT-1999; 99US-0161993.  
 29-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 647; DB 21; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-56;  
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 129 MRKAGVCFSEVPPDRKGMGVVDSYNNDDMKYAIRKLDATFPRNAFSAIYRVREYSR 188  
 |||||  
 Db 1 MRKAGVCFSEVPPDRKGMGVVDSYNNDDMKYAIRKLDATFPRNAFSAIYRVREYSR 60  
 |||||

QY 189 SVSRSPDDSKYSRSRSRGPGSCSYSSKSRSVSPARSISPRGRPLSRGRSLYSSVSRSGS 248  
 |||||  
 Db 61 SVSRSPDDSKYSRSRSRGPGSCSYSSKSRSVSPARSISPRGRPLSRGRSLYSSVSRSGS 120  
 |||||

QY 249 LIRAGDWM 256  
 |||||  
 Db 121 LIRAGDWM 128

RESULT 14  
 AAG30418  
 ID AAG30418 standard; Protein; 294 AA.  
 XX AAG30418;  
 AC AAG30418;  
 XX 17-OCT-2000 (first entry)  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 36362.  
 DE Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX EF1033405-A2.  
 PN 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-0301439.  
 PF 25-FEB-1999; 99US-0121825.  
 XX 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 08-APR-1999; 99US-0128234.  
 PR 16-APR-1999; 99US-0128714.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.

27-JUL-1999; 99US-0145919.  
 28-JUL-1999; 99US-0145951.  
 28-AUG-1999; 99US-0146386.  
 02-AUG-1999; 99US-0146388.  
 02-AUG-1999; 99US-0146389.  
 03-AUG-1999; 99US-0147038.  
 04-AUG-1999; 99US-0147204.  
 04-AUG-1999; 99US-0147302.  
 05-AUG-1999; 99US-0147192.  
 05-AUG-1999; 99US-0147260.  
 06-AUG-1999; 99US-0147303.  
 06-AUG-1999; 99US-0147416.  
 09-AUG-1999; 99US-0147493.  
 09-AUG-1999; 99US-0147935.  
 10-AUG-1999; 99US-0148171.  
 11-AUG-1999; 99US-0148319.  
 12-AUG-1999; 99US-0148341.  
 13-AUG-1999; 99US-0148565.  
 13-AUG-1999; 99US-0148684.  
 16-AUG-1999; 99US-0149368.  
 17-AUG-1999; 99US-0149175.  
 18-AUG-1999; 99US-0149426.  
 20-AUG-1999; 99US-0149722.  
 20-AUG-1999; 99US-0149723.  
 20-AUG-1999; 99US-0149929.  
 23-AUG-1999; 99US-0149902.  
 23-AUG-1999; 99US-0149930.  
 25-AUG-1999; 99US-0150566.  
 26-AUG-1999; 99US-0150884.  
 27-AUG-1999; 99US-0151065.  
 27-AUG-1999; 99US-0151066.  
 27-AUG-1999; 99US-0151080.  
 30-AUG-1999; 99US-0151303.  
 31-AUG-1999; 99US-0151438.  
 01-SEP-1999; 99US-0151930.  
 07-SEP-1999; 99US-0152363.  
 10-SEP-1999; 99US-0153070.  
 13-SEP-1999; 99US-0153758.  
 15-SEP-1999; 99US-0154018.  
 16-SEP-1999; 99US-0154039.  
 20-SEP-1999; 99US-0154779.  
 22-SEP-1999; 99US-0155139.  
 23-SEP-1999; 99US-0155486.  
 24-SEP-1999; 99US-0155659.  
 28-SEP-1999; 99US-0156458.  
 29-SEP-1999; 99US-0156596.  
 04-OCT-1999; 99US-0157117.  
 05-OCT-1999; 99US-0157753.  
 06-OCT-1999; 99US-0157865.  
 07-OCT-1999; 99US-0158029.  
 08-OCT-1999; 99US-0158232.  
 12-OCT-1999; 99US-0158369.  
 13-OCT-1999; 99US-0159293.  
 13-OCT-1999; 99US-0159294.  
 13-OCT-1999; 99US-0159295.  
 14-OCT-1999; 99US-0159329.  
 14-OCT-1999; 99US-0159330.  
 14-OCT-1999; 99US-0159331.  
 14-OCT-1999; 99US-0159637.  
 14-OCT-1999; 99US-0159638.  
 18-OCT-1999; 99US-0159584.  
 21-OCT-1999; 99US-0160741.  
 21-OCT-1999; 99US-0160767.  
 21-OCT-1999; 99US-0160768.  
 21-OCT-1999; 99US-0160770.  
 21-OCT-1999; 99US-0160814.  
 21-OCT-1999; 99US-0160815.  
 22-OCT-1999; 99US-0160980.  
 22-OCT-1999; 99US-0160981.  
 22-OCT-1999; 99US-0160989.  
 25-OCT-1999; 99US-0161404.  
 25-OCT-1999; 99US-0161405.  
 25-OCT-1999; 99US-0161406.

26-OCT-1999; 99US-0161359.  
 26-OCT-1999; 99US-0161360.  
 26-OCT-1999; 99US-0161361.  
 28-OCT-1999; 99US-0161920.  
 28-OCT-1999; 99US-0161922.  
 28-OCT-1999; 99US-0161993.  
 29-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 647; DB 21; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-56;  
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 129 MRKAGVCFSEVPPDRKMGSGVVDYNSYDDMKYAIRKLDATFPRNAFSAIYRVREYSR 188  
 |||||  
 Db 1 MRKAGVCFSEVPPDRKMGSGVVDYNSYDDMKYAIRKLDATFPRNAFSAIYRVREYSR 60  
 |||||

QY 189 SVSRSPDDSKYSRSRSGPCSCSYSSKSRSVSPARSISPRGRPLSRGRSLYSSVSRSGS 248  
 |||||  
 Db 61 SVSRSPDDSKYSRSRSGPCSCSYSSKSRSVSPARSISPRGRPLSRGRSLYSSVSRSGS 120  
 |||||

QY 249 LIRAGDWM 256  
 |||||  
 Db 121 LIRAGDWM 128  
 |||||

RESULT 14  
 AAG30418  
 ID AAG30418 standard; Protein; 294 AA.  
 XX AAG30418;  
 AC AAG30418;  
 XX  
 XX 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 36362.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW Hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX EF1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 PF  
 XX  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 08-APR-1999; 99US-0128234.  
 PR 16-APR-1999; 99US-0128714.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0141554.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148584.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 06-OCT-1999; 99US-0157753.  
PR 07-OCT-1999; 99US-0157865.  
PR 08-OCT-1999; 99US-0158029.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.  
 Query Match 44.6%; Score 646.5; DB 21; Length 294;  
 Best Local Similarity 69.7%; Pred. No. 2.1e-55;  
 Matches 124; Conservative 16; Mismatches 21; Indels 17; Gaps 2;  
 QY 1 MSSRWRTIYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPRPPGYAFVEFEDPRDADD 60  
 Db 1 MSSRSKRTIYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPRPPGYAFVEFEDPRDADD 60  
 QY 61 AIYGRDGYDFDGLRLRVEIAHGGRRFSPVDRYSSGSA-----SRAPSR 106  
 Db 61 AIYGRDGYDFDGLRLRVEIAHGGRRFSPVDRYSSGSA-----SRAPSR 117  
 QY 107 SDYRVLVTGLPPSASQDLKDHVRKAGDVCSEVFPDRKMGVVDYSDYDMKYAIR 164  
 Db 118 SEYRVVVSGLPSSASQDLKDHVRKAGDVCSEVFPDRKMGVVDYSDYDMKYAIR 175

RESULT 15  
 AAB11418  
 ID AAB11418 standard; Protein; 248 AA.  
 AC AAB11418;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human SP2/ASF protein.  
 XX  
 KW SR protein; splice-factor activity; plant; developmental behavior;  
 XX  
 OS flowering; crop plant; cereal; bean; rice; fruit.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200065059-A1.  
 XX  
 PD 02-NOV-2000.  
 XX  
 XX 20-APR-2000; 2000WO-AT00100.  
 XX  
 XX 23-APR-1999; 99AT-0000727.  
 XX  
 PA (OSTP ) OESTERR FORSCH SEIBERSDORF.  
 XX  
 XX Barta A, Lopato S, Kalyana M, Dörner S;  
 PI  
 XX WPI; 2000-687349/67.  
 DR

Novel proteins with splice-factor activity in plants, useful e.g. for  
 altering flowering time or development, and the nucleic acid that  
 encodes it -  
 XX  
 PS Disclosure; Fig 2; 67pp; German.  
 XX  
 XX This invention describes a novel protein (I) with splice-factor activity  
 in plants (I) modifies the choice of splice sites in many plant  
 CC pre-mRNAs. (I) (also the nucleic acid that encodes them and related  
 CC vectors or expression systems) are used: (i) to alter splice patterns in  
 CC plants, or their parts; (ii) to alter developmental behavior of plants;  
 CC and/or (iii) to delay flowering, particularly by at least 25% relative  
 CC to the wild type, especially in crop plants such as cereals, beans, rice  
 CC and fruit.  
 XX

SQ Sequence 248 AA;  
 Query Match 40.1%; Score 582.5; DB 21; Length 248;

Best Local Similarity 53.6%; Pred. No. 3.5e-49;  
 Matches 134; Conservative 28; Mismatches 47; Indels 41; Gaps 8;  
 QY 9 IYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPRPPGYAFVEFEDPRDADDALYGRDGY 68  
 Db 18 IYVGNLPGDIRKEVEDLFYKGPVVDIDLKIPRPPGYAFVEFEDPRDADDALYGRDGY 77  
 QY 69 DFDGCRRLRVEIAHGGRR-----RFSFSDRYSSSYSSASRAPSRSDYRV 111  
 Db 78 DYDGYRLRVEIPRSGRGTGCGGGGGGAPRGRYGP-----PSRRSENRV 123  
 QY 112 LVTGLPPSASQDLKDHVRKAGDVCSEVFPDRKMGVVDYSDYDMKYAIRKLDATF 171  
 Db 124 VVSLPFGSGWQDLKDHVRKAGDVCYADVVDG---TGVEFVRKEDMTYAVRKLDTKF 180  
 QY 172 R-NAFSSAYIRVREYERSVSRSPDDSKSYRSRSGRSPCSYSSKSRSPARS-SPR 229  
 Db 181 RSHEGETAYIRVKVDGPRSPSYGRSRRS-RSRSRSRRS---NSRSPSYSPRRSGSPR 236  
 QY 230 SRPL-SRPRS 238  
 Db 237 YSPRHSRPRS 246  
 Search completed: January 29, 2004, 02:50:13  
 Job time : 51 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:50:45 ; Search time 28 Seconds  
(without alignments)  
958.252 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRWRTIYVGNLPGDIRK.....RSKSRSRNSPVSFVIG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1159	79.9	237	2 E86223	hypothetical prote
2	865	59.6	303	2 F86158	alternative splici
3	839	57.8	303	2 S71185	splicing factor SF
4	646.5	44.6	294	2 T01307	alternative splici
5	597.5	41.2	243	2 T46221	PRE-MRNA SPLICING
6	582.5	40.1	248	2 A40040	alternative splici
7	517	35.6	221	2 S53075	splicing factor SR
8	498.5	34.4	201	2 C40040	alternative splici
9	497	34.3	201	2 S26404	alternative splici
10	497	34.3	292	2 B40040	alternative splici
11	435.5	30.0	344	2 S59043	splicing factor SR
12	432.5	29.8	350	2 A40459	nuclear phosphopro
13	431	29.7	474	2 A48133	pre-mRNA splicing
14	418.5	28.8	292	2 S53042	splicing factor SR
15	414	28.5	374	2 A37282	52K active chromatin
16	409	28.2	269	2 B47112	growth response pr
17	315.5	21.7	281	2 T26084	hypothetical prote
18	297	20.5	238	2 A57198	splicing factor, a
19	294	20.3	312	2 T26085	hypothetical prote
20	280	19.3	361	2 T42525	splicing factor-li
21	280	19.3	365	2 T37730	probable pre-mRNA
22	237.5	16.4	221	2 B42701	PR264 protein - ch
23	235	16.2	208	2 T16953	hypothetical prote
24	234.5	16.2	221	2 A42701	splicing factor SF
25	234	16.1	164	2 I54089	pre-mRNA splicing
26	234	16.1	164	2 S14016	X16 protein - mous
27	233	16.1	179	2 T34145	hypothetical prote
28	216	14.9	287	2 T50647	serine/arginine-ri
29	211	14.5	349	2 F85294	splicing factor At

splicing factor-li  
splicing factor SR  
splicing factor RS  
hypothetical prote  
splicing factor RS  
RNA-binding protei  
splicing factor 9G  
splicing factor RS  
myelin regulatory  
probable arginine/  
probable arginine/  
probable R32p22 sp  
hypothetical prote  
nucleolar protein  
probable RNA bindi  
mRNA splicing fact

30 208.5 14.4 302 2 T45890  
31 205 14.1 350 2 T05797  
32 204.5 14.1 284 2 T51304  
33 198.5 13.7 249 2 E84791  
34 193.5 13.3 271 2 T47978  
35 189 13.0 135 2 A46398  
36 188 13.0 200 2 T05112  
37 187 12.9 200 2 T52627  
38 186 12.8 274 2 A55335  
39 185.5 12.8 286 2 T09704  
40 181.5 12.5 250 2 A84905  
41 181 12.5 196 2 E84638  
42 173 11.9 196 2 T15917  
43 169.5 11.7 414 2 JN0866  
44 168.5 11.6 309 2 T47685  
45 163.5 11.3 275 2 T48696

## ALIGNMENTS

### RESULT 1

E86223  
hypothetical protein [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002

C/Accession: E86223

R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E86223

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-237 <STO>

A/Cross-references: GB:AE005172; NID:g3249109; PIDN:AAC24092.1; GSPDB:GN00141

C/Genetics:

C/Map position: 1

C/Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein

Query Match 79.9%; Score 1159; DB 2; Length 237;

Best Local Similarity 100.0%; Pred. No. 8.3e-87;

Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKGPVVDLKIIPRPPGYAFVEFEDPDADD 60

Db 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKGPVVDLKIIPRPPGYAFVEFEDPDADD 60

Qy 61 AIYGRDGYDPDGCRLVETAHGGRFSPVDRYSSYSASRAPSRSDYRVLTGLPPSA 120

Db 61 AIYGRDGYDPDGCRLVETAHGGRFSPVDRYSSYSASRAPSRSDYRVLTGLPPSA 120

Qy 121 SWDLDKDHMKACGVCFSEVFPDRKMGVVDYNSVDDMKYAIRKLDATFEFRNAPSAYI 180

Db 121 SWDLDKDHMKACGVCFSEVFPDRKMGVVDYNSVDDMKYAIRKLDATFEFRNAPSAYI 180

Qy 181 RVREYSRSVSRSPDCKSKYSRSRSGRGPSCSYSSKSR 218

Db 181 RVREYSRSVSRSPDCKSKYSRSRSGRGPSCSYSSKSR 218

### RESULT 2

F86158

alternative splicing factor SF2a [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002

```

Db      DB      61 AIHGRDGYDFDGHRLRYVELAHGRR-----SSDTRGSFNGGGRGGGRGGGSGRGPFSRRS 117
QY      QY      108 DYRVLVTGLPPSASWQDLKDHMRKAGDVCFSFYFPRKMGSGVVDYSNVDDMKYAIRKLD 167
DB      DB      118 EFRVLVTWLASSASWQDLKDHIAKGDCVFSQVYRARTTGVVDYTCYEDMKYALKLD 177
QY      QY      168 ATEFRNFPSSAYITRVREYRSVRSRSPDQSKSY-RSRSRSGFSCSYSKSRSVSPARSI 226
DB      DB      178 DTEFRNFPNGYVRVREYDGRKDSRSPSRGYSYKSRSRGRGSV---SRSRSRSRSRSR 234
QY      QY      227 SPRSRPISRSRSLYSVSRSGSLLRAGDWISQSRSKS--RSRSRNSPSPV 276
DB      DB      235 SPKAK--SSRSAPKSTSRSPGPR-----SKSRSPSRSRSRSRSPSPV 278

RESULT 4
T01307
alternative splicing factor ASF-2 homolog T14P8.21 - Arabidopsis thaliana
N:Alternate names: ribonucleoprotein
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 19-May-2000
C:Accession: T01307
R:Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana T14P8.
A:Reference number: 214290
A:Accession: T01307
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-294 <XAL>
A:Cross-references: EMBL:AF069298; NID:G3193282; PID:G3193304
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Intron: 29/3; 51/3; 76/3; 122/1; 137/3; 157/1; 176/1; 195/2
A:Notes: T14P8.21
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein repeat-containing proteins; ribonucleoprotein repeat-containing proteins
F:78-72/Domain: ribonucleoprotein repeat homology <RRM1>
F:121-186/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 44.6%; Score 646.5; DB 2; Length 294;
Best Local Similarity 69.7%; Pred. NO. 4.4e-45;
Matches 124; Conservative 16; Mismatches 21; Indels 17; Gaps 2

QY      QY      1 MSSRWRTIYVGNLPGDIRCEVEDLFYKGPVTDIDKIPRPPGYAFVEFEDPRDADD 60
DB      DB      1 MSSRSSTIYVGNLPGDIRREVEDLFSGKGVQDIDKIPRPPGYAFVEFEDADADD 60
QY      QY      61 AIYGRDGYDFDGCRLRVEIAHGGRFRFSPSYDRYSYSSA-----SRAPSR 106
DB      DB      61 AIYGRDGYDFDGHRLRYVELAHGRRSSHDA---RGSYSRGRGRGGGGRGGRGPFSRR 117
QY      QY      107 SDYRVLVTGLPPSASWQDLKDHMRKAGDVCFSFYFPRKMGSGVVDYSNVDDMKYAIR 164
DB      DB      118 SEYRVVWSGLPSSASWQDLKDHMRKGGEVCFSCVFRDGRGTTGIVDYSYEDMKYAVR 175

RESULT 5
T46221
PRE-MRNA SPLICING FACTOR SF2-like protein - Arabidopsis thaliana
N:Alternate names: protein T9C5.30
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T46221
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.
submitted to the Protein Sequence Database, December 1999
A:Reference number: 223026
A:Accession: T46221
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-243 <RIE>
A:Cross-references: EMBL:AL1132964

```





R.Roth, M.B.; Zahler, A.M.; Stolk, J.A.  
J. Cell Biol. 115, 587-596, 1991  
A>Title: A conserved family of nuclear phosphoproteins localized to sites of polymerase  
A:Reference number: A40459; MUID:92011300; PMID:1717489  
A:Accession: A40459  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-350 <ROT>  
A:Cross-references: EMBL:X58720; NID:g8496; PIDN:CAA41556.1; PID:g8497  
C:Genetics:  
A:Gene: SR55  
A:Cross-references: FlyBase:FBgn0004587  
A:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein  
C:Keywords: phosphoprotein; pre-mRNA splicing  
F:5-64/Domain: ribonucleoprotein repeat homology <RRM3>  
F:116-178/Domain: ribonucleoprotein repeat homology <RRM2>  
  
Query Match 29.8%; Score 432.5; DB 2; Length 350;  
Best Local Similarity 40.4%; Pred. No. 1.3e-27;  
Matches 124; Conservative 40; Mismatches 92; Indels 51; Gaps 9;  
  
Qy 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFDPDPRDADDAIYGRDGY 68  
Db 6 VVVGGLPYGVREDLIERFKYGRTRDILK-----NGYGFVEFEDYDADDAVYELNGK 60  
  
Qy 69 DFDGCLRLVEIAHGRFPSPVDYSSSY-----SASR-APSRSDYRVLV 113  
Db 61 ELGLGVVVEPARGARGG-NRDYDDRYGRRGGGGRYNEKSSRYGPPLRTEYRLIV 119  
  
Qy 114 TGLPSSAQDLKDHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKLDATPR- 172  
Db 120 ENLSRSVQDLKDMYRQAGEVYADAHKQRN-EGVVEFASLSMDKTAIKLDDTELNG 178  
  
Qy 173 -----NAFSSAIRVREYSRSPDPSKSYR---SRSRSGPSCSY 214  
Db 179 RRIHVEDRRGRSGGGGGRGRSSSRSSSRSSSRSSSRSSSRSSSRSSSRSG 238  
  
Qy 215 SKRSRSP-----ARSSPRSPRLSRSSLYSSVSRSGSLRAGDWMISQSR-----KSR 264  
Db 239 GRSKSKSPVKSRSRSGRSGRSDVSKSKSHSRTRSPKREDSRSTRYSKRSR 298  
  
Qy 265 SRSRSGNS 271  
Db 299 SRSRSGKS 305  
  
RESULT 13  
A48133  
pre-mRNA splicing SRP75 - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-1995 #sequence\_revision 30-Jun-1995 #text\_change 05-Nov-1999  
C:Accession: A48133  
R:Zahler, A.M.; Neugebauer, K.M.; Stolk, J.A.; Roth, M.B.  
Mol. Cell. Biol. 13, 4023-4028, 1993  
A>Title: Human SR proteins and isolation of a cDNA encoding SRP75.  
A:Reference number: A48133; MUID:93309435; PMID:8321209  
A:Accession: A48133  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-494 <ZAH>  
A:Cross-references: GB:L14076; NID:g307437; PIDN:AAA36649.1; PID:g307438  
A>Note: parts of this sequence were confirmed by peptide sequencing  
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein  
C:Keywords: phosphoprotein; pre-mRNA splicing  
F:3-62/Domain: ribonucleoprotein repeat homology <RRM3>  
F:105-167/Domain: ribonucleoprotein repeat homology <RRM2>  
  
Query Match 29.7%; Score 431; DB 2; Length 494;  
Best Local Similarity 40.1%; Pred. No. 2.7e-27;  
Matches 124; Conservative 39; Mismatches 92; Indels 54; Gaps 10;  
  
Qy 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFDPDPRDADDAIYGRDGY 68  
Db 6 VVVGGLPYGVREDLIERFKYGRTRDILK-----NGYGFVEFEDYDADDAVYELNGK 60

Db 4 VYIGRLSYQARERDVERFFKYGKILEVDLK-----NGYGFVEFDDLRDADDAVYELNGK 58  
Qy 69 DFDGCLRLVEIAHGRFPSPVDYSSSYASRA-----PSRSDYRVLVTLGLPSSASW 122  
Db 59 DLGGERVIVEHARGPRR-DGSYGSGRSGYGRSGRDKYGPPTREYRLIVENLSRCSW 117  
Qy 123 QDLKDHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKLDATPR----- 172  
Db 118 QDLKDYNRQAGEVYADAHKQRN-EGVIEFVSYDMKGALEKLDGTEVNGRKIRLVEDK 176  
Qy 173 --NAFSSAIRVREYSRSPVSR-----PDDSKYSRSGRSGPSCSY-SKRSRV 220  
Db 177 PGRRRRSYRSSH-SRSGRSPRSHKSRSGSGSKSHSKSRSGRSGRSGRSGR 235  
Qy 221 SPARSISPR-----SRPLSRSSLYSSVSRSGSLRAGDWI-----SOSRSK 262  
Db 236 SQGRSKSKKESRSPSKDKSRSSSHSAGKSRSKQDAEKEIQNDNVGPKPSRSPSRHK 295  
  
Qy 263 SRSRSGNS 271  
Db 296 SKSKSRSGS 304  
  
RESULT 14  
S59042  
splicing factor Srp40 - human  
C:Species: Homo sapiens (man)  
C>Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 28-May-1999  
C:Accession: S59042  
R:Streaton, G.R.; Caceres, J.F.; Mayeda, A.; Bell, M.V.; Flebanski, M.; Jackson, D.G.; B  
EMBO J. 14, 4336-4349, 1995  
A>Title: Identification and characterization of three members of the human SR family of  
A:Reference number: S59042; MUID:96016206; PMID:7556075  
A:Accession: S59042  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-272 <SCR>  
A:Cross-references: NID:g1049079; PIDN:AAA93070.1; PID:g1049080  
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein  
F:5-64/Domain: ribonucleoprotein repeat homology <RRM3>  
F:109-171/Domain: ribonucleoprotein repeat homology <RRM2>  
  
Query Match 28.8%; Score 418.5; DB 2; Length 272;  
Best Local Similarity 42.5%; Pred. No. 1.3e-26;  
Matches 121; Conservative 29; Mismatches 90; Indels 45; Gaps 10;  
  
Qy 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPPPGYAFVEFDPDPRDADDAIYGRDGY 68  
Db 6 VYIGRLNPAAREKDVVERFFKYGIRIDILK-----RGFGFVEFEDPRDADDAVYELDGK 60  
  
Qy 69 DFDGCLRLVEIAH-----GGRFSPSPVDYSSSYASR-----APSRSDYRVLV 115  
Db 61 EL--CSERVITIEHARARSGR-----GRGYSRDRSSRRPRNDRNAPPVVTENRLIVEN 114  
  
Qy 116 LPPSAGWQDLKDHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKLDATPR 175  
Db 115 LSRVSWQDLKDFWRQAGEVTFADAHKPLN-EGVVEFASYGDLKNAIEKLSKEI-NGR 172  
  
Qy 176 SSAYIRVREYSRSPVSRSPDPSKYSRSGRSGPSCSY-SKRSRSGPSPARSISPRSP 235  
Db 173 KIKLIGSKSHSRSGRSGRSTRS-SRSGRSGRSGRSGRSGRSGRSGRSGRSGR 222  
  
Qy 236 SRSLSVSRSG-----GSLLRAGDWISQSRSGRSGRSGRSGNS 271  
Db 223 SRSKSRSGRSPVPEKSKQKSGSSRSGSPASVDRQSRSGRSGRSGRSGRSGRSGR 267  
  
RESULT 15  
A37282  
52K active chromatin boundary protein - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 28-Feb-1992 #sequence\_revision 28-feb-1992 #text\_change 02-Jul-1998  
C:Accession: A37282

Search completed: January 29, 2004, 03:02:27  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 21:21:46 ; Search time 18 Seconds  
(without alignments)

728.914 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSRRNRRTIYVGNLFGDIRK.....RSKSRSRSRNSPVSFVSG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	865	59.6	303	1	SFR1_ARATH
2	582.5	40.1	247	1	SFR1_HUMAN
3	517	35.6	221	1	SFR9_HUMAN
4	440.5	30.4	344	1	SFR6_HUMAN
5	431	29.7	494	1	SFR4_HUMAN
6	426.5	29.4	489	1	SFR4_MOUSE
7	421	29.0	375	1	SFR5_DROME
8	418.5	28.8	272	1	SFR5_HUMAN
9	409	28.2	269	1	SFR5_RAT
10	374.5	25.8	270	1	SFR5_MOUSE
11	297	20.5	238	1	SFR2_HUMAN
12	280	19.3	365	1	SFR2_SCHPO
13	237.5	16.4	221	1	SFR2_CHICK
14	234.5	16.2	221	1	SFR2_MOUSE
15	234	16.1	164	1	SFR3_HUMAN
16	232	16.0	208	1	YSX2_CAEEL
17	226.5	15.6	221	1	SFR2_HUMAN
18	2210	14.5	356	1	RS41_ARATH
19	207.5	14.3	264	1	RS31_ARATH
20	205	14.1	350	1	RS40_ARATH
21	189	13.0	135	1	RBPI_DROME
22	173	11.9	196	1	SFR2_CAEEL
23	172	11.9	283	1	SFR2_MOUSE
24	169.5	11.7	414	1	NOP3_YEAST
25	164.5	11.3	275	1	SFR1_SCHPO
26	162	11.2	391	1	ROG_HUMAN
27	161.5	11.1	197	1	RX21_DROME
28	156.5	10.8	388	1	ROG_MOUSE
29	154.5	10.6	437	1	RUI7_HUMAN
30	153	10.5	429	1	HRBI_YEAST
31	152.5	10.5	378	1	RUI7_MOUSE
32	150.5	10.4	629	1	PAB2_ARATH
33	147	10.1	386	1	TIAL_MOUSE

34 146.5 10.1 289 1 TR2B\_HUMAN Q15815 homo sapien  
35 145.5 10.0 471 1 RUI7\_XENLA P09406 xenopus lae  
36 145 10.0 834 1 CWFV\_SCHPO Q9P639 schizosacch  
37 144.5 10.0 386 1 TIAL\_HUMAN P31483 homo sapien  
38 144 9.9 272 1 TRAZ\_DROVI O02008 drosophila  
39 142.5 9.8 632 1 PABP\_DROME P21187 drosophila  
40 141.5 9.8 289 1 ROCI\_ARATH Q9U044 arabidopsis  
41 140.5 9.7 273 1 ROCI\_NICSY Q08935 nicotiana s  
42 140.5 9.7 1089 1 Y555\_HUMAN Q9UKJ3 homo sapien  
43 139 9.6 264 1 TRAZ\_DROME P19018 drosophila  
44 138 9.5 282 1 ROC\_XENLA P19600 xenopus lae  
45 137.5 9.5 427 1 GBP2\_YEAST P25555 saccharomyc

## ALIGNMENTS

### RESULT 1

SFR1\_ARATH STANDARD; PRT; 303 AA.  
AC Q22315; Q22314; Q39201;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Pre-mRNA splicing factor SF2 (SRI protein).  
GN SF2 OR ATIG02840 OR F22D16.16 OR F22D16\_30.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95372342; PubMed=7644475;  
RA Lazar G., Schaal T., Maniatis T., Goodman H.M.;  
RT "Identification of a plant serine-arginine-rich protein similar to the  
mammalian splicing factor SF2/ASF.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:7672-7676(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Su C.-L., Schuler M.A.;  
RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA MEDLINE=21018719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,  
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
Dunn P., Etgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,  
Mitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana.";  
RL Nature 408:816-820(2000).  
CC -!- FUNCTION: CAN PROMOTE SPLICE SITE SELECTION IN VITRO PRESUMABLY BY  
ANTAGONIZING THE EFFECTS OF THE A1 HETEROGENEOUS NUCLEAR  
RIBONUCLEOPROTEIN. MAY HAVE AN ESSENTIAL FUNCTION DURING EARLY  
PLANT DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
Event-Alternative splicing; Named isoforms=2;

CC Name=1; Synonyms=SF2A;  
CC IsoId=022315-1; Sequence=Displayed;  
CC Name=2; Synonyms=SF2B;  
CC IsoId=022315-2; Sequence=VSP\_005859, VSP\_005860;  
CC TISSUE SPECIFICITY: Ubiquitous.  
CC PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS  
CC DOMAIN.  
CC SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.  
CC SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M98340; AA32856.1; -;  
CC EMBL; AF001035; AAB71385.1; -;  
CC EMBL; AF001035; AAB71386.1; -;  
CC EMBL; AC009525; AAF02881.1; -;  
CC PIR; F86158; F86158.  
CC PIR; S71185; S71185.  
CC HSP; P19339; 1SXL.  
CC InterPro; IPR000504; RNA\_rec\_mot.  
CC SMART; SM00076; rrm; 2.  
CC SMART; SM00360; RRM; 2.  
CC PROSITE; PS0102; RRM; 2.  
CC PROSITE; PS00030; RRM\_RNP\_1; FALSE NEG.  
CC RNA processing; mRNA splicing; Nuclear protein; RNA-binding; Repeat;  
CC Phosphorylation; Alternative splicing.  
CC DOMAIN 7 82 RNA-BINDING (RRM) 1.  
CC DOMAIN 96 112 GLY-RICH (HINGE REGION).  
CC DOMAIN 119 197 RNA-BINDING (RRM) 2.  
CC DOMAIN 198 273 ARG/SER-RICH (RS DOMAIN).  
CC DOMAIN 274 303 LYS/PRO/SER-RICH (PSK DOMAIN).  
CC VARSPIC 267 272 SRSRSR -> YGTYD (in isoform 2).  
CC VARSPIC 273 303 Missing (in isoform 2).  
CC CONFLICT 125 125 /FTid=VSP\_005860  
CC CONFLICT 127 127 G -> A (IN REF. 1).  
CC CONFLICT 139 140 P -> W (IN REF. 1).  
CC CONFLICT 303 303 MR -> IA (IN REF. 1).  
CC SEQUENCE 303 AA; 33729 MW; 6530F9CB628B8EFA CRC64;  
  
Query Match 59.6%; Score 865; DB 1; Length 303;  
Best Local Similarity 63.0%; Pred. No. 1.3e-60;  
Matches 184; Conservative 33; Mismatches 45; Indels 30; Gaps 7;  
  
Qy 1 MSSRWRTIYVGNLPGDIRKCEVEDLRYKGPVIVDIDLKIPPPPGVAFVEFEDPRADD 60  
Db 1 MSSRSRTYVGNLPGDIREREVEDLFSKGPVQVQIDLKVPFPFGVAFVEFDDARDAED 60  
  
Qy 61 AIYGRDGYDFGCLRLVRIAHGRRFSPVDRYSSYS- - - - -ASRAPSRSS 107  
Db 61 AIHGRDGYDFGHLRLVRLAAGRR- - -SSDTRGSFNGGGRGGRGDSGSGSPSRSS 117  
  
Qy 108 DYRLVLTGLPSPASQDLKDMRKAGDVCFSEVPFPRKMGSVYDYNDDMKYAIKRLD 167  
Db 118 EFRVLVLTGLPSPASQDLKDMRKAGDVCFSEVQVYDARGITGVVDYTCYEDMKYALKLD 177  
  
Qy 168 ATEPNASSYAIRVEYESVSRSPPDSKSY-RSRSRSGPSCSVSSKRSRSPARSI 226  
Db 178 DTEPNASSYVIRVEYDYSRKDSRSGRSRSGRSRSGV- - -SSRSRSRSGRSR 234  
  
Qy 227 SPRSRPLSRSLYSSVSRSGSLLRAGDWIQSRSKS- - -RSRSRNSGVPSPV 276  
Db 235 SPKAK- - -SRSPSPAKSTSRSPGPR- - - - -SKRSRSPSRSRSRSPPLSPV 278  
  
RESULT 2  
SFRL\_HUMAN

ID ID SFRL\_HUMAN STANDARD; PRT; 247 AA.  
AC Q07955; Q13809;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Splicing factor, arginine/serine-rich 1 (pre-mRNA splicing factor SF2,  
DE P33 subunit) [Alternative splicing factor ASF-1].  
GN SFPS1 OR SF2P33 OR ASF OR SF2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] \_  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 143-160 AND 166-174.  
RX MEDLINE=91309150; PubMed=1830244;  
RA Krainer A.R., Mayeda A., Kozak D., Binns G.;  
RT "Functional expression of cloned human splicing factor SF2: homology  
RT to RNA-binding proteins, U1 70K, and Drosophila splicing  
RT regulators";  
RL Cell 66:383-394 (1991).  
RN [2]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.  
RX MEDLINE=91309149; PubMed=1855257;  
RA Ge H., Zuo P., Manley J.L.;  
RT "Primary structure of the human splicing factor ASF reveals  
RT similarities with Drosophila regulators";  
RL Cell 66:373-382 (1991).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM ASF-1).  
RC TISSUE=Placenta;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [4]  
RP SEQUENCE OF 122-139.  
RX MEDLINE=92249775; PubMed=1577277;  
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;  
RT "SR proteins: a conserved family of pre-mRNA splicing factors";  
RL Genes Dev. 6:837-847 (1992).  
RN [5]  
RP INTERACTIONS IN SPLICEOSOME ASSEMBLY.  
RX MEDLINE=94084782; PubMed=8261509;  
RA Wu J.Y., Maniatis T.;  
RT "Specific interactions between proteins implicated in splice site  
RT selection and regulated alternative splicing";  
RL Cell 75:1061-1070 (1993).  
RN [6]  
RP FUNCTION IN RECRUITMENT OF U1-70K TO PRE-MRNA.  
RX MEDLINE=94187841; PubMed=8139654;  
RA Kohetz J.D., Jamison S.F., Will C.L., Zuo P., Luhrmann R.,  
RA Garcia-Blanco M.A., Manley J.L.;  
RT "Protein-protein interactions and 5'-splice-site recognition in  
RT mammalian mRNA precursors";  
RL Nature 368:119-124 (1994).  
RN [7]





DB 180 RSHEGETAYIVKVDGPRSPSYGSRSSRS-RSRSSRSRS---NSRSRSYSPRRSGSPR 235

QY 230 SRPL-SRGRS 238

DB 236 YSPRRSGSRSS 245

RESULT 3

SFR9 HUMAN

ID SFR9\_HUMAN STANDARD; PRT; 221 AA.

AC Q13242;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Splicing factor, arginine/serine-rich 9 (Pre-mRNA splicing factor

DE SRP30C).

GN SFRS9 OR SRP30C.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon;

RA MEDLINE=96016206; PubMed=7556075;

RA Sreaton G.R., Caceres J.F., Mayeda A., Bell M.V., Plebanski M.,

RA Jackson D.G., Bell J.I., Krainer A.R.;

RT "Identification and characterization of three members of the human SR

RT family of pre-mRNA splicing factors.";

RL EMBO J. 14:4336-4349(1995).

RN [2]

RP SEQUENCE FROM N.A.

RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Southam L., Boriraj V.,

RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,

RA Lathrop M., Cox R.D., Bell G.I.;

RT "Transcription map of the 5CM region surrounding the hepatocyte

RT nuclear factor-1a/MODY3 gene on chromosome 12.";

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Murphy L.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: PLAYS A ROLE IN CONSTITUTIVE SPLICING AND CAN MODULATE

CC THE SELECTION OF ALTERNATIVE SPLICE SITES.

CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE KIDNEY, LUNG

CC AND SPLEEN, INTERMEDIATE LEVELS IN THE LIVER AND SMALL INTESTINE,

CC AND LOW LEVELS IN THE BRAIN, HEART AND SKELETAL MUSCLE.

CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS

CC DOMAIN (BY SIMILARITY).

CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.

CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DB EMBL; U30825; AAA93059.1; -

DB EMBL; U87279; AAD00826.1; -

DB EMBL; U87277; AAD00826.1; JOINED.

DB EMBL; U87278; AAD00826.1; JOINED.

DB EMBL; AL021546; CAA16498.1; -

DB PIR; S59075; S59075.

DB Genew; HGNC:10791; SFRS9.

DB GK; Q13242; -

DB MIM; 601943; -

DB GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.

DB GO; GO:0006376; P:mRNA splice site selection; TAS.

DB InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam: PF00076; RRM; 2.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS0102; RRM; 2.

DR PROSITE; PS00030; RRM RNP 1; FALSE\_NEG.

KW Nuclear protein; RNA-binding; mRNA splicing; Repeat; Phosphorylation.

FT DOMAIN 14 89 RNA-BINDING (RRM) 1.

FT DOMAIN 91 100 GLY-RICH (HINGE REGION).

FT DOMAIN 111 187 RNA-BINDING (RRM) 2.

FT DOMAIN 188 200 ARG/SER-RICH (RS DOMAIN).

SQ SEQUENCE 221 AA; 25542 MW; 1EE7BD8601CD80C0 CRC64;

Query Match 35.6%; Score 517; DB 1; Length 221;

Best Local Similarity 52.5%; Pred. No. 1.3e-33;

Matches 115; Conservative 30; Mismatches 38; Indels 36; Gaps 8;

QY 9 IVVGNLPGDIRKCEVEDLFYKYGPIVDIDLK-----IPPPPGYAFVFEFDPDADDAIY 63

DB 16 IVVGNLPTDVREKLEDFYKYGRIETELKVRHGLVP-----FAFVAFEDPRDAEDAIY 70

QY 64 GRDGVDFPGCLRLVEI--AHGGRFPSPVDYSSYSASRAPSRSDYRVLTGLPPSAS 121

DB 71 GRNGYDYGQCLRLVEFPRTYGGRGWPRGGR-----NGPPTRRSDRFLVSLGPPSGS 123

QY 122 WDLKDHMRKAGDVCFSVFPDRKMGVDPVDSYVDDMKYAIRKLDATERS-NAFSSAYI 180

DB 124 WDLKDHMRKAGDVCYADVQKDGVM---VEYLKEDMEYAIRKLDITKFRSHEGETSYI 180

QY 181 RVREYESRSVSRSPDDSKSY---RSRSRSGSPSCSYSSK 216

DB 181 RV-----YPERSTSYGSRSGSRGSDSPYQSR 209

RESULT 4

SFR6 HUMAN

ID SFR6\_HUMAN STANDARD; PRT; 344 AA.

AC Q13247; Q13244; Q13245; Q96J06; Q9UJB8; Q9Y3N7;

DT 15-DEC-1998 (Rel. 37, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Splicing factor, arginine/serine-rich 6 (Pre-mRNA splicing factor

DE SRP55).

GN SFR56 OR SRP55.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE=Colon;

RX MEDLINE=96016206; PubMed=7556075;

RA Sreaton G.R., Caceres J.F., Mayeda A., Bell M.V., Plebanski M.,

RA Jackson D.G., Bell J.I., Krainer A.R.;

RT "Identification and characterization of three members of the human SR

RT family of pre-mRNA splicing factors.";

RL EMBO J. 14:4336-4349(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.D.,

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Connor R.R.,

RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhali P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA Levaeslaibo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConnachie I.J., McLay K., McMurray A.A.,

RA Milne S.A., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871 (2001).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM SRP55-2).  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM SRP55-1).  
RN [4]  
RC TISSUE=Placenta;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [5]  
RP SEQUENCE OF 21-27 AND 47-55.  
RX MEDLINE=92249775; PubMed=1577277;  
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;  
RT "SR proteins: a conserved family of pre-mRNA splicing factors.";  
RL Genes Dev. 6:837-847 (1992)  
CC -!- FUNCTION: PLAYS A ROLE IN CONSTITUTIVE SPLICING AND CAN MODULATE  
CC THE SELECTION OF ALTERNATIVE SPLICING SITES.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=SRP55-1;  
CC -isoId=Q13247-1; Sequences=Displayed;  
CC Name=SRP55-2;  
CC -isoId=Q13247-2; Sequences=VSP\_005869, VSP\_005870;  
CC Name=SRP55-3;  
CC -isoId=Q13247-3; Sequences=VSP\_005871;  
CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS  
CC DOMAIN (BY SIMILARITY).  
CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.  
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U30863; AAA93073.1; -;  
DR EMBL; U30828; AAA93071.1; -;  
DR EMBL; U30829; AAA93072.1; -;

DR EMBL; AL031681; CAB43960.1; -;  
DR EMBL; AL031681; CAB43961.1; -;  
DR EMBL; BC006832; AAH06832.1; -;  
DR PIR; S59043; S59043;  
DR Genew; HGNC:10788; SFRS6.  
DR GK; Q13247; -;  
DR MIM; 601944; -;  
DR GO; GO:0008248; Fipre-mRNA splicing factor activity; TAS.  
DR GO; GO:0006376; P.mRNA splice site selection; TAS.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS0102; RRM; 2.  
DR PROSITE; PS00030; RRM\_RNP\_1; FALSE NEG.  
KW Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;  
KW Repeat; Phosphorylation.  
FT DOMAIN 1 72 RNA-BINDING (RRM) 1.  
FT DOMAIN 87 90 GLY-RICH (HINGE REGION).  
FT DOMAIN 110 183 RNA-BINDING (RRM) 2.  
FT DOMAIN 184 343 ARG/SER-RICH (RS DOMAIN).  
FT VARSPLIC 86 135 SGGGGSSRRSTGDKIGPFPVTEYELIVENLSSRSQWDL  
FT KDFMQAGE -> MTNGAEAVSTEAKMTAPDFWFLHTLC  
FT DPCPMTLWLTLPAMTAAFC (in isoform  
FT SRP55-2).  
FT /FTId=VSP\_005869.  
FT Missing (in isoform SRP55-2).  
FT /FTId=VSP\_005870.  
FT RVSPPPKATSRSSRSRSRSRSRSRD -> LKLG  
FT RFMSQQTSLYSLASSC (in isoform SRP55-3).  
FT /FTId=VSP\_005871.  
FT /FTId=VSP\_005871.  
FT R -> H (IN REF. 1).  
FT CONFLICT 64 64  
SQ SEQUENCE 344 AA; 39586 MW; 72305506CE948B94 CRC64;  
  
Query Match 30.4%; Score 440.5; DB 1; Length 344;  
Best Local Similarity 39.1%; Pred. No. 2.2e-27;  
Matches 124; Conservative 44; Mismatches 88; Indels 61; Gaps 12;  
  
Qy 9 IYVGNLPGDIRKCEVEDLFYKYPVIDIDILKIPRPFGYAFVEFFEDPRADDAIYGRDGY 68  
Db 4 VYIGRLSYNVREKDIQRFSGYGRLLVDLK-----NGYGFVEFFEDSRDADDAVYELNGK 58  
  
Qy 69 DFDGCLRLVEIAHGRRFPSPVDYS-----SSYSASRAPSR-----RSDYRLVLT 114  
Db 59 ELGGERVIVEHARGPRR---DRDGYSGRSGGGYSSRRSTGSDKYGPPVRYEYLIVE 115  
Qy 115 GLPPSASWDLKDHRKAGDVCSEVPDRKGMGVVDYSDNYDMKYAIRKLDATF--- 171  
Db 116 NLSRCSWDLKDFMQAGEVYADAHKRTN-EGVIEFESYSDMKRALDKLDGTINGR 174  
Qy 172 -----RNAFSSAYIRVR-----EYESRSVSRSPDSDSKSYRSRSGRP 209  
Db 175 NIRLIEDKPTSHRRSYSGSRSSRRSSRRSSRRSSRRSSRSISKRS-RRSRSKGR 233  
Qy 210 SCYSS--KSRVSPARSISPR---SRPLSRSLYSSVSRSGSLIRA-----GDWISQ 258  
Db 234 SRSRSGKRSKRSKSKPKSDRGSHSRSDKDEYK-SRSRSRSGRPENKGDIKSK 292  
Qy 259 SRSKSRSGRSPNSPVSP 275  
Db 293 SRSRSGRSPNSPLVPP 309  
  
RESULT 5  
SFRA\_HUMAN STANDARD; PRT; 494 AA.  
AC Q08170; Q9BUA4; Q9UEB5;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Splicing factor, arginine/serine-rich 4 (Pre-mRNA splicing factor  
DE SRP75) (SRP0011LB).  
GN SFRS4 OR SRP75.  
OS Homo sapiens (Human).

GO	GO:0005634; C:nucleus; TAS.
DR	GO:0008248; F:pre-mRNA splicing factor activity; TAS.
DR	GO:0006371; P:mRNA splicing; TAS.
GO	InterPro; IPR000504; RNA_rec_mot.
DR	Pfam; PF00076; rrm; 2.
DR	SMART; SM00360; RRM; 2.
DR	PROSITE; PS50102; RRM; 2.
DR	PROSITE; PS00300; RRM_RNP_1; FALSE NEG.
DR	mRNA processing; mRNA splicing; Nuclear protein; RNA-binding;
KW	Repeat; Phosphorylation.
FT	DOMAIN 2 72
FT	DOMAIN 72 97
FT	DOMAIN 104 177
FT	DOMAIN 179 494
FT	CONFLICT 35 35
FT	CONFLICT 253 253
FT	CONFLICT 318 322
FT	CONFLICT 338 338
FT	CONFLICT 356 356
FT	CONFLICT 436 438
FT	SEQUENCE 494 AA; 56678 MW; 5BBAB917C218C20A CRC64;
SQ	

Query Match                  29.7%; Score 431; DB 1; Length 494;  
Best Local Similarity      40.1%; Pred. No. 1.9e-26;  
Matches 124; Conservative    39; Mismatches 92; Indels    54; Gaps    10;

QY	9	IYVGNLPGDIRKCEVEDLFYKYGPVIDIDLKIPRPPGVAFFVEFDPDADDAIYGRDGY	68
DB	4	VVIGLSYCAREDDVERFFKGKYLEVDLK-----NGVGFEFDLRDADDAAVELNGK	58
QY	69	PDFGRLRVETIAHGRRSPSVDRISSSYASRA-----PERRDYRVLTGLPPSAW	122
DB	59	DLCGBRVTEHARGPRR-DGSYGSGRGYGRRSGRDKYGPPTRTYEYLIVENLSSRCSW	117
QY	123	ODLDKHMKAQGVCFSEVPFRPKGMGVVDYSNDDMKYAIRKLDATEPR-----	172
DB	118	QDLKKMFMQAAGVTYADAHKKN-EGVIEFVSYDMKALKELDTENGKRILVEDX	176
QY	173	--NAFSSAYIRVREYSRSVRS-----PDDSKSYESRSRSGPCSCYS-SKSRSV	220
DB	177	PGSRRRSYSRGRSH-SRSRSRSHRKSRSSGSSKSHSKSRSGSRSGSRKSRSR	235
QY	221	SPARISPR-----SRPLSRGRSLYSSVSRSGLLRAGDWI-----SQSRSK	262
DB	236	SQSRSRCKEKSRSPEKESRSRSHSAGKRSKSDQAEKIQNNDNVGKPXSFSRHK	295
QY	263	SESRSRNS 271	
DB	296	SKSKSRSS 304	

RESULT 6  
SFZ4\_MOUSE  
ID\_SFZ4\_MOUSE STANDARD; PRT; 489 AA.  
AC Q8V597; Q9JJC3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Splicing factor, arginine/serine-rich 4.  
GN SFRS4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

[illegible]

QY	172	----	--RNAFS-----		----	---SAYIRVRE-----	YSR
Dd	170	IRLVEDKPGSRRRSYGRSHRSRSHSRSHSRSGSSKSHSKGRSRSGSHSR		----	---	----	229
QY	189	SVSRSPPDDSKSVYRSRSGSPSCSYSGKSRSVP-		----	---	----	225
Dd	230	SKVSRSGQRGRSKKEKSRGPFKDNRSGRSPDKSRKSKDHAEDKLQNDSAGAKS		----	---	----	289
QY	226	TSP-----	-----RSRPISRSRSYYSVSRSGSLLRAGDWISOSRSKRSR		----	----	266
Dd	290	HPSFHDHSKRSRGORRAEEERRRVSRARSGSEKRSQEKSLLK-----	-----SRSRSGRSR		----	----	344
QY	267	SRS	269		----	----	
Dd	345	SRS	347		----	----	

RESULT 7  
SR35\_DROME

ID	SR55_DROME	STANDARD;	PRT;	375 AA.
AC	P26686; Q24252;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Serine-arginine protein 55 (SRP55) (Enhancer of deformed) (52-kDa Serine-arginine protein) (B52 protein).			
GN	B(DFP) OR SR55 OR R355 OR B52.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OX	Ephydroidea; Drosophilidae; Drosophila.			
NCBI_TaxID=7227;				
[1]	SEQUENCE FROM N.A., AND SEQUENCE OF 1-14; 125-131 AND 136-147.			
RN	STRAIN=CL; TISSUE=Embryo;			
RC	MEDLINE=92011900; PubMed=1717489;			
RX	Roth M.B., Zahler A.M., Stolk J.A.;			
RT	"A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription.";			
RL	J. Cell Biol. 115:587-596(1991).			
[2]	SEQUENCE FROM N.A. (ISOFORM LONG).			
RN	TISSUE=Embryo;			
RC	MEDLINE=91357476; PubMed=1885003;			
RX	Champlin D.T., Frasch M., Saumweber H., Lis J.T.;			
RA	"Characterization of a Drosophila protein associated with boundaries of transcriptionally active chromatin.";			
RT	Genes Dev. 5:1611-1621(1991).			
RL	[3]			
RN	CHARACTERIZATION.			
RP	MEDLINE=92159037; PubMed=1741384;			
RX	Mayeda A., Zahler A.M., Krainer A.R., Roth M.B.;			
RA	"Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 89:1301-1304 (1992).			
RL	[4]			
RN	FUNCTION			
RP	MEDLINE=95021280; PubMed=7935465;			
RX	Ring H.-Z., Lis J.T.;			
RA	"The SR protein B52/SRP55 is essential for Drosophila development.";			
RT	Mol. Cell. Biol. 14:7499-7506(1994).			
RL	-1- FUNCTION: ESSENTIAL FOR DEVELOPMENT. MAY HAVE A CRITICAL ROLE IN SPLICING OR IN CONTROLLING ALTERNATIVE SPLICE SITE USE OF AT LEAST SOME PRE-MRNA IN VIVO. NOT REQUIRED FOR ALL SPLICING. MAY PLAY A GENERAL ROLE IN THE CONDENSATION OR DECONDENSATION OF CHROMATIN.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR; ASSOCIATED WITH BOUNDARIES OF TRANSCRIPTIONALLY ACTIVE CHROMATIN.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=2;			
CC	Name=Long;			
CC	IsoId=P26686-1; Sequence=Displayed;			
CC	Name=Short;			
CC	IsoId=P26686-2; Sequence=VSP 005878;			



```

RESULT 9
ID SFR5_RAT STANDARD; PRT; 269 AA.
ID SFR5_RAT STANDARD; PRT; 269 AA.
Q09167; Q05335;
01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
DE Splicing factor, arginine/serine-rich 5 (Pre-mRNA splicing factor
DE SRP40) (Insulin-induced growth response protein CL-4) (Delayed-early
DE protein HRS).
GN SFRS5 OR CL-4 OR HRS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxId=10116;
[1]
RN
RP SEQUENCE FROM N.A.
RC
RC TISSUE=Liver;
RX MEDLINE=93315501; PubMed=7686911;
RA Diamond R.H., Du K., Lee V.N., Mohr K.L., Haber B.A., Tewari D.S.,
RA Taub R.;
RT "Novel delayed-early and highly insulin-induced growth response
RT genes. Identification of HRS, a potential regulator of alternative
RT pre-mRNA splicing.";
RL J. Biol. Chem. 268:15185-15192(1993).
[2]
RN
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC
RC TISSUE=Thymus;
RX MEDLINE=98094279; PubMed=9434190;
RA Du K., Taub R.;
RT "Alternative splicing and structure of the human and mouse
RT SFRS5/HRS/Srp40 genes.";
RL Gene 204:243-249(1997).
[3]
RN
RP SEQUENCE OF 185-269 FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=94164020; PubMed=8161377;
RA Hamil K.G., Hall S.H.;
RT "Cloning of rat Sertoli cell follicle-stimulating hormone primary
RT response complementary deoxyribonucleic acid: regulation of TSC-22
RT gene expression.";
RL Endocrinology 134:1205-1212(1994).
CC -! FUNCTION: MAY BE REQUIRED FOR PROGRESSION THROUGH G1 AND ENTRY
CC INTO S PHASE OF CELL GROWTH. MAY PLAY A REGULATORY ROLE IN PRE-
CC MRNA SPLICING. AUTOREGULATES ITS OWN EXPRESSION. PLAYS A ROLE IN
CC CONSTITUTIVE SPLICING AND CAN MODULATE THE SELECTION OF
CC ALTERNATIVE SPLICING SITES (BY SIMILARITY).
CC -! FUNCTION: COULD PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
CC DIFFERENTIATION IN THE SPLEEN AND THYMUS.
CC -! ALTERNATIVE PRODUCTS:
CC Event=alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q09167-1; Sequence=Displayed;
CC Name=2; Synonyms=HRR-LF;
CC IsoId=Q09167-2; Sequence=VSP_005867, VSP_005868;
CC -! TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN AND THYMUS.
CC -! INDUCTION: BY INSULIN AND HEPATECTOMY.
CC -! PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN (BY SIMILARITY).
CC -! SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -! SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; LI3635; AAA62286.1; -
CC DR
CC EMBL; AF020683; AAA471864.1; -
CC DR
CC EMBL; LI3267; AAA42316.1; -
CC DR

```





GO:	GO:0008248;	P:pre-mRNA splicing factor activity; TAS.
GO:	GO:0006371;	P:mRNA splicing; TAS.
InterPro:	IPR000504;	RNA_rec_mot.
InterPro:	IPR001878;	Znf_CCHC.
Pfam:	PF00076;	zrm; 1.
Pfam:	PF00098;	zf-CCHC; 1.
PRINTS:	PRO00339;	C2HCZNFINGER.
SMART:	SM00360;	RRM; 1.
SMART:	SM00343;	Znf_C2HC; 1.
PROSITE:	PS50102;	RRM; 1.
PROSITE:	PS00030;	RRM_RNP_1; FALSE_NEG.
PROSITE:	PS50158;	ZF_CCHC; 1.
KW	Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;	
KW	Phosphorylation; Repeat; Zinc-finger.	
FT	DOMAIN	11 84
FT	ZN_FING	104 120
FT	DOMAIN	121 238
FT	DOMAIN	153 227
FT	REPEAT	153 160
FT	REPEAT	161 169
FT	REPEAT	169 176
FT	REPEAT	177 184
FT	REPEAT	211 218
FT	REPEAT	219 226
FT	VARSPLIC	130 135
FT	VARSPLIC	136 238
FT	VARSPLIC	130 132
FT	VARSPLIC	133 238
FT	SEQUENCE	238 AA; 27366 MW; 49136754D9630853 CRC64;
QY	Query Match	20.58; Score 297; DB 1; Length 238;
Db	Best Local Similarity	37.1%; Pred. No. 2.2e-16; Matches 102; Conservative 22; Mismatches 73; Indels 78; Gaps 12
QY	3 SRWR- ----TIYGNLPGDIRKEVEDIFYKYGPIVDLKIIPRPFGYAFVEFEDPRD	57
Db	2 SRYGYGGEIKVYVGNLTGAGKGELEAFSYGPLRTV--WIAPPGFAPVEFEDPRD	59
QY	58 ADDAIYRGDGYDPGCGRLRVEIAHGRRSPSPVDYSSSYASRAPSRSDYRVLTGLP	117
Db	60 AEDAARGLDGKVIKCGSRVRVELSTG---MPPRSRF-----DRPFARRP-----	99
QY	118 PSASQDLKQHRKAGDVCSEVFDPKCGSWGVDYSDMKYAIRKLDATEFENAFSS	177
Db	100 -----FDPNRCYECGE-----KG-----HYAYDCHRY-----	123
QY	178 AYIVREYERYSRSPDDSKSY-RGRSRSGPSCSYSSKSRSPARSISPRGRPLSR	236
Db	124 ---RRRSRSPSRSHSRGRYRSRSGSRG-----RRSRASPRSS---RSISLRSS	171
QY	237 RSLYSSYSRSGSLRAGDWTISQSRSKSRSRSSNS	271
Db	172 RSALRSRSGSI--KGSRYFQSPSRSRSRSSIS	204
RESULT 12		
SRP2_SCHPO		
ID	SRP2_SCHPO	STANDARD; PRT; 365 AA.
AC	P78614;	
DT	28-FEB-2003	(Rel. 41, Created)
DT	28-FEB-2003	(Rel. 41, Last sequence update)
DT	28-FEB-2003	(Rel. 41, Last annotation update)
DE	Pre-mRNA splicing factor srp2.	
GN	SRP2 OR SPAC16.02C.	
OS	Schizosaccharomyces pombe (fission yeast).	
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;	
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;	
OC	Schizosaccharomycos	
NCBI	TaxID=4896;	

[1]  
PP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=9608180; PubMed=9421507;  
RA Gross T., Richter K., Mierke C., Luetzelberger M., Kaeuffer N.F.;  
RT "Identification and characterization of sp1, a gene of fission yeast  
RT encoding a RNA binding domain and a RS domain typical of SR splicing  
RT factors.";  
RL Nucleic Acids Res. 26:505-511(1998).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=972;  
RX MEDLINE=2184401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgueros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins S., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voickart G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Snpakovski G.V., Usery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
[3]  
RN SEQUENCE OF 5-365 FROM N.A.  
RP STRAIN=PR745;  
RX MEDLINE=98162722; PubMed=9501991;  
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;  
RT "Identification of open reading frames in Schizosaccharomyces pombe  
RT cDNAs.";  
RL DNA Res. 4:363-369(1997).  
[4]  
CC -!- FUNCTION: Has a role in pre-mRNA splicing where it is involved in  
CC spliceosome assembly (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- PTM: Extensively phosphorylated on serine residues in the RS  
CC domain (By similarity).  
CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.  
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF012278; AAC39357.1; -;  
CC EMBL; AL121745; CAB57400.1; -;  
CC EMBL; D89163; BAA13825.1; -;  
CC PIR; T37730; T37730;  
CC PIR; T42525; T42525;  
CC GeneDB SPombe; SPAC16.02c; -;  
CC InterPro; IPR000504; RNA\_rec\_mot.  
CC Pfam; PF00076; rrm; 2;  
CC SMART; SM00360; RRM; 2;  
CC PROSITE; PS50102; RRM; 2;

DR PROSITE; PS00030; RRM RNP 1; FALSE NEG.  
KW mRNA processing; mRNA\_splicing; Nuclear protein; RNA-binding;  
KW Repeat; Phosphorylation.  
FT DOMAIN 6 69 RNA-BINDING (RRM) 1.  
FT DOMAIN 74 91 GLY-RICH (HINGE REGION).  
FT DOMAIN 100 166 RNA-BINDING (RRM) 2.  
FT DOMAIN 182 300 ARG/SER-RICH (RS DOMAIN).  
FT SEQUENCE 365 AA; 42566 MW; 67E988573A736691 CRC64;  
Query Match 19.3%; Score 280; DB 1; Length 365;  
Best Local Similarity 31.6%; Pred. No. 8e-15;  
Matches 85; Conservative 38; Mismatches 114; Indels 32; Gaps 7;  
QY 9 IYVGNLPDGRKCEVEDLFYKYGIVDIDILKPPPCYAFVEPRDADDAIYGRDGY 68  
DB 6 LFGVRIIPQATREDMDFKYGQIQDLCKLM-----NGFGEVEVEDARDADI VNDFOGK 60  
QY 69 DFDGRLRVEIAGHGRRSPSDVRYSSYSASAPSRSDYRVLVTGLTPPSASWQDLKDH 128  
DB 61 EFMGSRIVVEFARGERR---RRENFRESAASKYPRPRTGRLIVENLSEDSVQDLKDV 117  
QY 129 MRKAGDVCFSEVFPDRKMGSGVDYSNYDDMKYAIRKLD-----ATEFRNAPSAYIR 181  
DB 118 MRKAGEPIFTDAHRENPAG-AGVVEFSTEEDVRNALTSLNGEVIKQAVTLREDPDAA NEP 176  
QY 182 VREYRSRVSPPDDSKYRSRSRSPSCYSYSSKRSVSPARSISPRSRPLSRSLYS 241  
DB 177 LPEVPSRFRSPSPARRRYDDYR-RGGDYRRDAYRFGDRDDERYAPRGEYRRNRDEY- 234  
QY 242 SVRSRGSLLRAGDMISQSRSKRSRSRSD 270  
DB 235 -----RRG-----GRDYRRNSRSD 249  
RESULT 13  
SPR2\_CHICK STANDARD; PRT; 221 AA.  
AC P30352;  
DT 01-APR-1993 (Rel. 25; Created)  
DT 01-APR-1993 (Rel. 25; Last sequence update)  
DT 28-FEB-2003 (Rel. 41; Last annotation update)  
DE Splicing factor, arginine/serine-rich 2 (Splicing factor SC35) (SC-35)  
DE (Splicing component, 35 kDa) (PR264 protein).  
GN SPRS2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thymus;  
RX MEDLINE=92212859; PubMed=1557353;  
RA Vellard M., Sureau A., Soret J., Martinerie C., Perbal B.;  
RT "A potential splicing factor is encoded by the opposite strand of the  
RT trans-spliced c-myc exon.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2511-2515(1992).  
CC -!- FUNCTION: NECESSARY FOR THE SPLICING OF PRE-MENA. IT IS REQUIRED  
CC FOR FORMATION OF THE EARLIEST ATP-DEPENDENT SPLICING COMPLEX AND  
CC INTERACTS WITH SPLICEOSOMAL COMPONENTS BOUND TO BOTH THE 5' AND 3'  
CC SPICE SITES DURING SPLICEOSOME ASSEMBLY. IT ALSO IS REQUIRED FOR  
CC ATP-DEPENDENT INTERACTIONS OF BOTH U1 AND U2 SNRNPS WITH PRE-  
CC MENA.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS  
CC DOMAIN (BY SIMILARITY).  
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way



RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-37 AND 54-64.  
RC SPECIES=Human;  
RX MEDLINE=92249775; PubMed=1577277;  
RA Zahler A.M., Lane W.S., Stolk J.A., Roth M.B.;  
RT "SR proteins: a conserved family of pre-mRNA splicing factors.";  
RL Genes Dev. 6:837-847(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RA Liu W.L., Wang M., Tang D., Rodgers G.;  
RT "Identification and characterization of novel full-length cDNAs  
RT differentially expressed in hematopoietic lineages";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Placenta;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.  
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Igoe L.P., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse;  
RX MEDLINE=91232908; PubMed=2030943;  
RA Ayane M., Preuss U., Koehler G., Nielsen P.J.;  
RT "A differentially expressed murine RNA encoding a protein with  
RT similarities to two types of nucleic acid binding motifs.";  
RL Nucleic Acids Res. 19:1273-1278(1991).  
RN [5]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC SPECIES=Mouse; STRAIN=MRL; TISSUE=Lymphoid;  
RX MEDLINE=97299661; PubMed=9154810;  
RA Jumaa H., Guenet J.-L., Nielsen P.J.;  
RT "Regulated expression and RNA processing of transcripts from the  
RT Sp20 splicing factor gene during the cell cycle.";  
RL Mol. Cell. Biol. 17:3116-3124(1997).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse; STRAIN=C57BL/6J;  
RX MEDLINE=21085560; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Blake J., Boffelli D., Furuno M., Anon H., Baldarelli R., Bash G.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weiz C., Whittaker C., Wilming L.,

RA Myshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC !- FUNCTION: MAY BE INVOLVED IN RNA PROCESSING IN RELATION WITH  
CC CELLULAR PROLIFERATION AND/OR MATURATION.  
CC !- SUBCELLULAR LOCATION: Nuclear.  
CC !- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=P23152-1; Sequence=D:played;  
CC Name=Short;  
CC IsoId=P23152-2; Sequence=VSP\_005861, VSP\_005862;  
CC Note=Has been shown to exist only in mouse so far;  
CC !- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THYMUS AND PRE-B CELL  
CC LINES; HIGH, IN TESTIS, BRAIN AND SPLEEN; VERY LOW IN HEART AND  
CC NOT DETECTABLE IN LIVER AND KIDNEY.  
CC !- INDUCTION: BY SERUM; IN A TISSUE CULTURE.  
CC !- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS  
CC DOMAIN (BY SIMILARITY).  
CC !- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC !- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
CC EMBL; L10838; AAA36648.1; -;  
CC EMBL; AF107405; AAD44523.1; -;  
CC EMBL; BC000914; AAH00914.1; -;  
CC EMBL; X53824; CAA37821.1; -;  
CC EMBL; X91556; CAA62844.1; -;  
CC EMBL; X91556; CAA62845.1; -;  
CC EMBL; AK011657; BAB27762.1; -;  
CC PIR; I54089; I54089.  
CC PIR; S14016; S14016.  
CC HSP; F11340; IGVJ.  
CC Genew; HGNC:10785; SFRS3.  
CC MIM; 603364; -;  
CC MIM; P23152; -;  
CC MGD; MGI:98285; Sfrs3.  
CC GO; GO:0003723; F:RNA binding activity; TAS.  
CC InterPro; IPR000504; RNA\_rec\_mot.  
CC Pfam; PF00076; rtm; 1.  
CC SMART; SM00360; RRM; 1.  
CC PROSITE; PS0102; RRM; 1.  
CC PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
CC Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;  
CC Phosphorylation; Repeat.  
CC FT DOMAIN 10 83 RNA-BINDING (RRM).  
FT DOMAIN 86 164 ARG/SER-RICH (RS DOMAIN).  
FT DOMAIN 119 164 2 X APPROXIMATE REPEATS, BASIC.  
FT REPEAT 119 133 B-1.  
FT REPEAT 149 164 B-2.  
FT VARSPLIC 115 124 SPFRSFRSRS -> VTMSLLTL (in isoform  
FT Short).  
FT FTId=VSP\_005861.  
FT Missing (in isoform Short).  
FT /FTId=VSP\_005862.  
SQ SEQUENCE 164 AA; 19329 MW; 02F0A5EE33FF28A0 CRC64;  
Query Match 16.1%; Score 234; DB 1; Length 164;  
Best Local Similarity 30.2%; Pred. No. 1.2e-11;  
Matches 79; Conservative 24; Mismatches 47; Indels 112; Gaps 9;  
QY 9 YVGNLPGDIRKCEVEDLFYKYGVIVDLKIPRPPGVAVEFEDPRDADDAIYGRDGY 68  
Db 12 YVGNLGNNGNKTLELAFGYGGLRSV--VWANNPGFAVEFEDPRDADDAVRELDGR 69

QY	69	DFDGCRLRVEIAHGGRFPSPVDYSSYSASRAPSRSDYRVLVTGLPPSASWQDLKDH	128
Db	70	TLCGCRVRVELSNGEKR-----SRNGP-----PPSMGRRPRDDY	104
QY	129	MRKAGDVCFSEVFPDRKMGSGVVDYKNYDDMKYAIRKLDATFERNAPSSAIRVREYESR	189
Db	105	RRSP-----PPRR-----	113
QY	189	SVSRSPDDSKSVSRSRSGPSCSVSKSRVSPARSISPRSRPLSRSLYSSVSRSGS	248
Db	114	---RSP-----RRSFER-----SRSLSRDRR---RERSLSRERNHKP-----	147
QY	249	LIRAGDWISQSRKSRSRSRN	270
Db	148	-----SRSFERSRSRSRN	161

Search completed: January 29, 2004, 02:50:43  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model.

```
Run on: January 29, 2004, 02:55:41 ; Search time 53 Seconds
        (without alignments)
        1358.428 Million cell updates/sec
```

Title: US-10-014-927-19  
 Perfect score: 1451  
 Sequence: 1 MSSRRNRITYVGNLPGDIRK.....RSKSRSRSRNSPVSPIVG 279

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

```
Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23; \*

- ```

1: sp_archea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_plant.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID     | Description        |
|------------|--------|-------------|--------|----|--------|--------------------|
| 1          | 1377.5 | 94.9        | 268    | 10 | Q9XFR5 | Q9Xfr5 arabidopsis |
| 2          | 1343   | 92.6        | 256    | 10 | Q9XFR6 | Q9Xfr6 arabidopsis |
| 3          | 1159   | 79.9        | 237    | 10 | Q8O496 | Q8O496 arabidopsis |
| 4          | 1032.5 | 71.2        | 207    | 10 | Q949S8 | Q949s8 arabidopsis |
| 5          | 863    | 59.5        | 307    | 10 | Q9SP11 | Q9Sp11 arabidopsis |
| 6          | 847    | 58.4        | 285    | 10 | Q8L7P1 | Q8L7p1 arabidopsis |
| 7          | 845    | 58.2        | 270    | 10 | Q9SP14 | Q9Sp14 arabidopsis |
| 8          | 845    | 58.2        | 276    | 10 | Q9SP15 | Q9Sp15 arabidopsis |
| 9          | 845    | 58.2        | 289    | 10 | Q9SP12 | Q9Sp12 arabidopsis |
| 10         | 827.5  | 57.0        | 380    | 10 | Q8H453 | Q8H453 cryza sativ |
| 11         | 806    | 55.5        | 295    | 10 | Q9CA06 | Q9Ca06 arabidopsis |
| 12         | 748    | 51.6        | 250    | 10 | Q8RZS6 | Q8Rzs6 cryza sativ |
| 13         | 745    | 51.3        | 261    | 10 | Q9SP13 | Q9Sp13 arabidopsis |
| 14         | 649.5  | 44.8        | 178    | 10 | Q8GX50 | Q8GX50 arabidopsis |
| 15         | 646.5  | 44.6        | 294    | 10 | Q8I290 | Q8I290 arabidopsis |
| 16         | 537.5  | 41.2        | 243    | 10 | Q9SC13 | Q9Sc13 arabidopsis |

## ALIGNMENTS

## RESULT 1

```

ID Q9XFR5 PRELIMINARY; PRT; 268 AA.
AC Q9XFR5;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE SF2/ASF-like splicing modulator Srp30 (Putative SF2/ASF splicing
DE modulator Srp30).
DE
DE SRF30 OR ATIG09140.
GN Arabidopsis thaliana (Mouse-ear cross).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosida
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99234087; PubMed=10215626;
RA Lopato S., Kalyva M., Dorner S., Kobayashi R., Krainer A.R., Barta
RT "atsrp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana
RT regulates splicing of specific plant genes.";
RL Genes Dev. 13:987-1001(1999).
RN [2]
RN SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.
RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA Ecker J.R., Theologis A.,
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (SEP-2002) to the EMBL/Genbank/DDBJ databases.
DR EMBL; AJ131214; CAB42557.1; -
DR EMBL; AY150486; AAM13011.1; -
DR HSSP; P08579; 1A2N.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SMO0360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 268 AA; 30385 MW; 73BDC35334A8F9AC4 CRC64;

```

|                           |                                                                        |                                                               |
|---------------------------|------------------------------------------------------------------------|---------------------------------------------------------------|
| Query Match               |                                                                        | 94.9%; Score 1377.5; DB 10; Length 268;                       |
| Best Local Similarity     |                                                                        | 96.1%; Pred. No. 1.4e-114;                                    |
| Matches 268; Conservative |                                                                        | 0; Mismatches 0; Indels 11; Gaps 1;                           |
| Qy                        | 1                                                                      | MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60 |
| Db                        | 1                                                                      | MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60 |
| Qy                        | 61                                                                     | AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120   |
| Db                        | 61                                                                     | AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120   |
| Qy                        | 121                                                                    | SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSAYI 180 |
| Db                        | 121                                                                    | SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSAYI 180 |
| Qy                        | 181                                                                    | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSRPLSRSRSLY 240  |
| Db                        | 181                                                                    | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSRPLSRSRSLY 240  |
| Qy                        | 241                                                                    | SSVSRSGSLLRAGDWISQSRKSRSRSGRSPVSPVISG 279                     |
| Db                        | 241                                                                    | SSVSR-----SQSRKSRSRSGRSPVSPVISG 268                           |
| RESULT 2                  |                                                                        |                                                               |
| Q9XFR6                    |                                                                        |                                                               |
| ID                        | Q9XFR6                                                                 | PRELIMINARY; PRT; 256 AA.                                     |
| AC                        | Q9XFR6                                                                 |                                                               |
| DT                        | 01-NOV-1999                                                            | (TrEMBLrel. 12, Created)                                      |
| DT                        | 01-NOV-1999                                                            | (TrEMBLrel. 12, Last sequence update)                         |
| DT                        | 01-OCT-2002                                                            | (TrEMBLrel. 22, Last annotation update)                       |
| DE                        | SP2/ASF-like splicing modulator Srp30, variant 1.                      |                                                               |
| GN                        | Srp30.                                                                 |                                                               |
| OS                        | Arabidopsis thaliana (Mouse-ear cress).                                |                                                               |
| OC                        | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;     |                                                               |
| OC                        | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  |                                                               |
| OC                        | eurosid II; Brassicales; Brassicaceae; Arabidopsis.                    |                                                               |
| OX                        | NCBI_TaxID=3702;                                                       |                                                               |
| RN                        | [1]                                                                    |                                                               |
| RP                        | SEQUENCE FROM N.A.                                                     |                                                               |
| RX                        | MEDLINE=99234087; PubMed=10215626;                                     |                                                               |
| RA                        | Lopato S., Kalyna M., Dorner S., Kobayashi R., Kraimer A.R., Barta A.; |                                                               |
| RT                        | "atSRp30, one of two SP2/ASF-like proteins from Arabidopsis thaliana,  |                                                               |
| RT                        | regulates splicing of specific plant genes."                           |                                                               |
| RL                        | Genes Dev. 13:987-1001(1999).                                          |                                                               |
| DR                        | EMBL; AJ131214; CAB42558.1; -.                                         |                                                               |
| DR                        | HSSP; P08579; 1A9N.                                                    |                                                               |
| DR                        | InterPro; IPR000504; RNA_rec_mot.                                      |                                                               |
| DR                        | Pfam; PF00076; rtm; 2.                                                 |                                                               |
| DR                        | SMART; SM00360; RRM; 2.                                                |                                                               |
| DR                        | PROSITE; PS0102; RRM; 2.                                               |                                                               |
| SQ                        | SEQUENCE 256 AA; 29113 MW; 834FE60B6E8FC25E CRC64;                     |                                                               |
| Query Match               |                                                                        | 92.6%; Score 1343; DB 10; Length 256;                         |
| Best Local Similarity     |                                                                        | 100.0%; Pred. No. 1.6e-111;                                   |
| Matches 256; Conservative |                                                                        | 0; Mismatches 0; Indels 0; Gaps 0;                            |
| Qy                        | 1                                                                      | MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60 |
| Db                        | 1                                                                      | MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60 |
| Qy                        | 61                                                                     | AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120   |
| Db                        | 61                                                                     | AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120   |
| Qy                        | 121                                                                    | SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSAYI 180 |
| Db                        | 121                                                                    | SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSAYI 180 |
| Qy                        | 181                                                                    | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSRPLSRSRSLY 240  |
| Db                        | 181                                                                    | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSRPLSRSRSLY 240  |

|                           |                                                                       |                                                               |
|---------------------------|-----------------------------------------------------------------------|---------------------------------------------------------------|
| Db                        | 181                                                                   | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRSVSPARSISPRSRPLSRSRSLY 240  |
| Qy                        | 241                                                                   | SSVSRSGSLLRAGDWI 256                                          |
| Db                        | 241                                                                   | SSVSRSGSLLRAGDWI 256                                          |
| RESULT 3                  |                                                                       |                                                               |
| O80496                    |                                                                       |                                                               |
| ID                        | O80496                                                                | PRELIMINARY; PRT; 237 AA.                                     |
| AC                        | O80496                                                                |                                                               |
| DT                        | 01-NOV-1998                                                           | (TrEMBLrel. 08, Created)                                      |
| DT                        | 01-NOV-1998                                                           | (TrEMBLrel. 08, Last sequence update)                         |
| DT                        | 01-OCT-2002                                                           | (TrEMBLrel. 22, Last annotation update)                       |
| DE                        | Tl2M4.19 protein.                                                     |                                                               |
| GN                        | Tl2M4.19                                                              |                                                               |
| OS                        | Arabidopsis thaliana (Mouse-ear cress).                               |                                                               |
| OC                        | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;    |                                                               |
| OC                        | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; |                                                               |
| OC                        | eurosid II; Brassicales; Brassicaceae; Arabidopsis.                   |                                                               |
| OX                        | NCBI_TaxID=3702;                                                      |                                                               |
| RN                        | [1]                                                                   |                                                               |
| RP                        | SEQUENCE FROM N.A.                                                    |                                                               |
| RC                        | STRAIN=ev. Columbia;                                                  |                                                               |
| RA                        | Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Oji, O., Kwan, A., |                                                               |
| RA                        | Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B., |                                                               |
| RA                        | Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,   |                                                               |
| RA                        | Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,            |                                                               |
| RA                        | Theologis A.;                                                         |                                                               |
| RT                        | "Arabidopsis thaliana chromosome 1 BAC Tl2M4 sequence, complete       |                                                               |
| RT                        | sequence."                                                            |                                                               |
| RL                        | Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.               |                                                               |
| RN                        | [2]                                                                   |                                                               |
| RP                        | SEQUENCE FROM N.A.                                                    |                                                               |
| RC                        | STRAIN=ev. Columbia;                                                  |                                                               |
| RA                        | Theologis A.;                                                         |                                                               |
| RL                        | Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.               |                                                               |
| DR                        | EMBL; AC003114; AAC24092.1; -.                                        |                                                               |
| DR                        | HSSP; P08579; 1A9N.                                                   |                                                               |
| DR                        | InterPro; IPR000504; RNA_rec_mot.                                     |                                                               |
| DR                        | Pfam; PF00076; rtm; 2.                                                |                                                               |
| DR                        | SMART; SM00360; RRM; 2.                                               |                                                               |
| DR                        | PROSITE; PS0102; RRM; 2.                                              |                                                               |
| SQ                        | SEQUENCE 237 AA; 27438 MW; AF63E0FF1E274F9E CRC64;                    |                                                               |
| Query Match               |                                                                       | 79.9%; Score 1159; DB 10; Length 237;                         |
| Best Local Similarity     |                                                                       | 100.0%; Pred. No. 3.4e-95;                                    |
| Matches 218; Conservative |                                                                       | 0; Mismatches 0; Indels 0; Gaps 0;                            |
| Qy                        | 1                                                                     | MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60 |
| Db                        | 1                                                                     | MSSRWNRITVYGNLPGDIRKCEVEDLFYKGPVIDIDLKIPRPPGYAFVEFEDPRDADD 60 |
| Qy                        | 61                                                                    | AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120   |
| Db                        | 61                                                                    | AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDYSSYSASRAPRRSDYRVLTGLPPSA 120   |
| Qy                        | 121                                                                   | SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSAYI 180 |
| Db                        | 121                                                                   | SWQDLKDHMRKAGDVCFSEVPFDRKMGSGVVDYSDYDMKYAIRKLDATFRNAPSAYI 180 |
| Qy                        | 181                                                                   | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRS 218                       |
| Db                        | 181                                                                   | RVREYESRSVSRSPDDSKSYRSRSGPSCSYSSKRS 218                       |
| RESULT 4                  |                                                                       |                                                               |
| Q949S8                    |                                                                       |                                                               |
| ID                        | Q949S8                                                                | PRELIMINARY; PRT; 207 AA.                                     |
| AC                        | Q949S8                                                                |                                                               |
| DT                        | 01-DEC-2001                                                           | (TrEMBLrel. 19, Created)                                      |
| DT                        | 01-OCT-2002                                                           | (TrEMBLrel. 22, Last sequence update)                         |
| DT                        | 01-MAR-2003                                                           | (TrEMBLrel. 23, Last annotation update)                       |

```
DE Putative SF2/ASF splicing modulator Srp30 protein (Fragment).
GN AT1G09140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050912; AAK93589.2; -.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
FT NON TER
SQ SEQUENCE 207 AA; 23336 MW; 76B74CEC8FD4772 CRC64;

Query Match 71.2%; Score 1032.5; DB 10; Length 207;
Best Local Similarity 94.0%; Pred. No. 5.3e-84;
Matches 205; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 62 IYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPSSAS 121
Db 1 IYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPSSAS 60
QY 122 WQDLKDHMRKAGDVCFSEVFPDRKMGSGVDPVSNYDDMKYAIRKLDATFENAFSSA 181
Db 61 WQDLKDHMRKAGDVCFSEVFPDRKMGSGVDPVSNYDDMKYAIRKLDATFENAFSSA 120
QY 182 VREYESRSVSPDCKSYRSRSGRSPSCSYSSKRSVSPARISPRSRPLSRSLYS 241
Db 121 VREYESRSVSPDCKSYRSRSGRSPSCSYSSKRSVSPARISPRSRPLSRSLYS 180
QY 242 SVSRSGSLLRAGDWISQSRSSKRSRSPVSPVSG 279
Db 181 SVSR-----SQSRSSKRSRSPVSPVSG 207

RESULT 5
Q9SP11 PRELIMINARY; PRT; 307 AA.
AC Q9SP11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Splicing factor SRI.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Landberg;
RA Lazar G., Goodman H.M.;
RT "The Arabidopsis splicing factor SRI is regulated by alternative
RT splicing.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF173640; AAD52609.1; -.
DR HSPF; P19339; ISXL.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.

SQ SEQUENCE 207 AA; 23336 MW; 76B74CEC8FD4772 CRC64;

Query Match 59.5%; Score 863; DB 10; Length 307;
Best Local Similarity 62.2%; Pred. No. 1.1e-88;
Matches 184; Conservative 33; Mismatches 45; Indels 34; Gaps 7;

QY 1 MSSRWNRRTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDDAD 60
Db 1 MSSRSRTIYVGNLPGDIREREVEDLFYKYGPIVDIDLVPPFPFYAFVEFDDAD 60
QY 61 AIYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPSSAS 103
Db 61 AIHGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPSSAS 117
QY 104 SRSDYRLVTGLPSSASQDLKDHMRKAGDVCFSEVFPDRKMGSGVDPVSNYDDMKY 163
Db 118 SRSEFRVLVTGLPSSASQDLKDHMRKAGDVCFSEVFPDRKMGSGVDPVSNYDDMKY 177
QY 164 RKLDATFENAFSSAIVREYESRSVSPDCKSYRSRSGRSPSCSYSSKRSVSP 222
Db 178 KKLDDTEFRNAPSNGYRVREYDNRKDSRSPGRSGRSGRSGV---SRSRSR 234
QY 223 ARSISPRSPRLSRSLYSVSRSGSLLRAGDWISQSRSSKRSRSPVSPV 276
Db 235 SRSRSPKAK--SRSRSPAKSTSRSPGPR-----SKSRSPSRSRSGRSGRSG 282

RESULT 6
Q8L7P1 PRELIMINARY; PRT; 285 AA.
AC Q8L7P1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE SF2/ASF-like splicing modulator Srp30, putative.
GN AT1G02840.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV128338; AAM91541.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 285 AA; 31967 MW; 85FC1FCAD24C9BF6 CRC64;

Query Match 58.4%; Score 847; DB 10; Length 285;
Best Local Similarity 61.9%; Pred. No. 2.7e-67;
Matches 177; Conservative 32; Mismatches 41; Indels 36; Gaps 6;

QY 1 MSSRWNRRTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDDAD 60
Db 1 MSSRSRTIYVGNLPGDIREREVEDLFYKYGPIVDIDLVPPFPFYAFVEFDDAD 60
QY 61 AIYGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPSSAS 107
Db 61 AIHGRDGYDFDGLRLVEIAHGRFSPSVDRYSYSSASRAPSRSDYRLVTGLPSSAS 117
QY 108 DYRLVTGLPSSASQDLKDHMRKAGDVCFSEVFPDRKMGSGVDPVSNYDDMKY 167
Db 118 EFRVLVTGLPSSASQDLKDHMRKAGDVCFSEVFPDRKMGSGVDPVSNYDDMKY 177
```



QY 168 ATERNAFSSAYITRVREYVSRSRDDSKSY-RSRSRSGPSCSYSSKSRSSVSPARS 226  
DB 178 DTEFRNFAFNGYVRVREYDSKXSGRSGRSKSRSGRSV---SRSRSRSGRSR 234  
QY 227 SPSRSP--SRSRSXSSVRSRSGLLRAGDWISQSRKSRSGRSRSGNSP 272  
DB 235 SPKAK--SSRSRSPAKTSRSPG-----PRSKSRSPSP 264

RESULT 7  
Q9SPI4  
ID AC Q9SPI4 PRELIMINARY; PRT; 270 AA.  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE Splicing factor SR1A.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lazar G., Goddman H.M.;  
RT "The Arabidopsis splicing factor SR1 is regulated by alternative  
splicing."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF173640; AAD52613.1; -  
DR HSSP; P19339; 1SXL.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 270 AA; 30147 MW; 0C9890D8C6A9E562 CRC64;

Query Match 58.2%; Score 845; DB 10; Length 270;  
Best Local Similarity 61.0%; Pred. No. 3.7e-67;  
Matches 177; Conservative 32; Mismatches 41; Indels 40; Gaps 6;

QY 1 MSSRWNTIYVGNLPGDIRKEVEDLFYKGPVVDLKIIPRPPGYAFVEFDPDADD 60  
DB 1 MSSRSRTVYVGNLPGDIREREVEDLFYKGPVVDLKVPRPPGYAFVEFDARDAED 60  
QY 61 AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDRYSSYS-----ASRAP 103  
DB 61 AIHGRDGYDFDGCRLRVELAHGGR---SSDTRGSGFNGGGRGGRGGRGDSGRGP 117  
QY 104 SRSRSDYRVLTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSNDDMKYAI 163  
DB 118 SRSSEFRVLVTGLFSSASWQDLKDHMRKGGDVCFQYVRDAGTTGVVDYTCYEDMKYAL 177  
QY 164 RKLDATERNAFSSAYITRVREYVSRSRSPDDSKSY-RSRSRSGPSCSYSSKSRSSVSP 222  
DB 178 KKLDDTEFRNFAFNGYVRVREYDSKXSGRSGRSKSRSGRSV---SRSRSRSGR 234  
QY 223 ARSISPRPLSRSGSYSSVRSRSGLLRAGDWISQSRKSRSGRSRSGNSP 272  
DB 235 SRSRSPKAK--SSRSRSPAKTSRSPG-----PRSKSRSPSP 268

RESULT 8  
Q9SPI5  
ID AC Q9SPI5 PRELIMINARY; PRT; 276 AA.  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE Splicing factor SR1C.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lazar G., Goddman H.M.;  
RT "The Arabidopsis splicing factor SR1 is regulated by alternative  
splicing."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF173640; AAD52611.1; -  
DR HSSP; P19339; 1SXL.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 276 AA; 30894 MW; A91150EB7D425C98 CRC64;

Query Match 58.2%; Score 845; DB 10; Length 276;  
Best Local Similarity 61.0%; Pred. No. 3.9e-67;  
Matches 177; Conservative 32; Mismatches 41; Indels 40; Gaps 6;

QY 1 MSSRWNTIYVGNLPGDIRKEVEDLFYKGPVVDLKIIPRPPGYAFVEFDPDADD 60  
DB 1 MSSRSRTVYVGNLPGDIREREVEDLFYKGPVVDLKVPRPPGYAFVEFDARDAED 60  
QY 61 AIYGRDGYDFDGCRLRVEIAHGGRRFSPVDRYSSYS-----ASRAP 103  
DB 61 AIHGRDGYDFDGCRLRVELAHGGR---SSDTRGSGFNGGGRGGRGGRGDSGRGP 117  
QY 104 SRSRSDYRVLTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKMGSGVVDYSNDDMKYAI 163  
DB 118 SRSSEFRVLVTGLFSSASWQDLKDHMRKGGDVCFQYVRDAGTTGVVDYTCYEDMKYAL 177  
QY 164 RKLDATERNAFSSAYITRVREYVSRSRSPDDSKSY-RSRSRSGPSCSYSSKSRSSVSP 222  
DB 178 KKLDDTEFRNFAFNGYVRVREYDSKXSGRSGRSKSRSGRSV---SRSRSRSGR 234  
QY 223 ARSISPRPLSRSGSYSSVRSRSGLLRAGDWISQSRKSRSGRSRSGNSP 272  
DB 235 SRSRSPKAK--SSRSRSPAKTSRSPG-----PRSKSRSPSP 268

RESULT 9  
Q9SPI2  
ID AC Q9SPI2 PRELIMINARY; PRT; 289 AA.  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE Splicing factor SR1B.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lazar G., Goddman H.M.;  
RT "The Arabidopsis splicing factor SR1 is regulated by alternative  
splicing."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF173640; AAD52610.1; -  
DR HSSP; P19339; 1SXL.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 289 AA; 32352 MW; F12C9F798A2235E8 CRC64;

Query Match 58.2%; Score 845; DB 10; Length 289;  
Best Local Similarity 61.0%; Pred. No. 4.1e-67;  
Matches 177; Conservative 32; Mismatches 41; Indels 40; Gaps 6;

```
QY 1 MSRWRTIYVGNLPGDIRKCEVDLFYKYGPVIVDIDLKIPRPPGYAFVFEFDPDADD 60
D 1 MSRSRTIYVGNLPGDIREREVEDLFYKYGVQIDLVKPPRPPGYAFVFEFDPDADD 60
QY 61 AIYGRDGYDFDGLRLVEIAHGRFPSPVDRYSSYS-----ASRAP 103
D 61 AIHGRDGYDFDGLRLVELAHGRR---SSDTRGSFNGGGRGGGRDGRDGGSRGP 117
QY 104 SRSRSDYVLTGLPSSASWQDLKHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAI 163
D 118 SRSRSEFVLTGLPSSASWQDLKHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAL 177
QY 164 RKLDATEFRNAPSAIYRVEYSRSPDSDSKY-RSRSRSGPSCSYSSKSRSP 222
D 178 KLLDTEFRNAPSAIYRVEYSRSPDSDSKY-RSRSRSGPSCSYSSKSRSP 234
QY 223 ARSISPRSPFLSRSLYSVSRSGSLLRAGDWISQSRKSRSRSP 272
D 235 SRSRSPKAK-SRSRSPAKTSRSPG-----FRSKSRSP 268

RESULT 10
QBH453
ID Q8H453 PRELIMINARY; PRT; 380 AA.
AC Q8H453;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE Putative pre-mRNA splicing factor SF2 (SR1 protein).
GN P0470D12.43.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0470D12.43";
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004300; BAC16007.1; -.
SQ SEQUENCE 380 AA; 42656 MW; C0B591F23B597022 CRC64;

Query Match 57.0%; Score 827.5; DB 10; Length 380;
Best Local Similarity 65.0%; Pred. No. 2.2e-65;
Matches 180; Conservative 23; Mismatches 49; Indels 25; Gaps 6;

QY 1 MSRWRTIYVGNLPGDIRKCEVDLFYKYGPVIVDIDLKIPRPPGYAFVFEFDPDADD 60
D 67 MSRWRTIYVGNLPGDIREREVEDLFYKYGPVIVDIDLKIPRPPGYAFVFEFDPDADD 126
QY 61 AIYGRDGYDFDGLRLVEIAHGRFPSPVDRYSSYSASR--APSRSDYVLTGLPP 118
D 127 AIRGRDGYDFDGLRLVELAHGRRGNS--FNSGGGGRGGVSRTEYRVLVTGLPS 183
QY 119 SASWQDLKHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYARKLDATEFRNAPSA 178
D 184 SASWQDLKHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYARKLDSBKNAPSKA 243
QY 179 YRVRVEY---SRSVSRSPDSDSKYRSRSGPSCSYSSKSRSPARSIPRSP 235
D 244 YRVRVEYDGRSRYSRSP-----RSRSGRSGYRSRSPRSGGKSPKGRSRSP 296
QY 236 SRSLSVSRSGSLLRAGDWISQSRKSRSRSP 272
D 297 SRSRSP-----SKGRSPSRSPARSQSP 323

RESULT 11
Q9CA06
PRELIMINARY; PRT; 295 AA.
AC Q9CA06;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Putative pre-mRNA splicing factor, 53460-55514.
GN TIG12.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haase B.J., Wu D.,
Baird R., Rensing C.M., Koo H., Fujii C.Y., Utterback T.R.,
Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 1 BAC TIG12 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012329; AGS2185.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 295 AA; 33070 MW; B257767C2876B9CB CRC64;

Query Match 55.5%; Score 806; DB 10; Length 295;
Best Local Similarity 61.8%; Pred. No. 1.3e-63;
Matches 183; Conservative 28; Mismatches 53; Indels 32; Gaps 10;

QY 1 MSRWRTIYVGNLPGDIRKCEVDLFYKYGPVIVDIDLKIPRPPGYAFVFEFDPDADD 60
D 1 MSRFRSRIYVGNLPGDIREREVEDLFYKYGRIVDIELKVPFRPPCYCFVEFEHSDAED 60
QY 61 AIYGRDGYDFDGLRLVEIAHGRFPSPVDRY-----SSSY-----SASRSPSR 106
D 61 AIKRGDGYNDGCLRLVELAHGRRGSS--DRGGYGGGGSGYGGGGGGGSGARFGVSRH 119
QY 107 SDYRVLVTGLPSSASWQDLKHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKL 166
D 120 SEFRVIVRGLPSSASWQDLKHMRKAGVCFSEVPDRKMGSGVVDYNDYDMKYAIRKL 179
QY 167 DATEFRNAPSAIYRVEYR-YSRVSRSPPDSDSKS-YRSRSGRSGPSCSY--SKSRSP 221
D 180 DTEFRNAPSAIYRVEYR-YSRVSRSPPDSDSKS-YRSRSGRSGPSCSY--SKSRSP 239
QY 222 PARSIPRSP-----SRPSRSLYSSVSRSGSLLRAGDWISQSRKSRSRSP 272
D 240 PRKDLSPRSLYSSVSRSGSLLRAGDWISQSRKSRSRSP-----MSRSGRSGRSP 288

RESULT 12
QBZS6
ID QBZS6 PRELIMINARY; PRT; 250 AA.
AC QBZS6;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DE Putative pre-mRNA splicing factor SF2.
GN B1012D10.26.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone: B1012D10.26";
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
```

DR EMBL; AP003535; BAB90350.1; -  
DR Gramene; Q8R2S6; -  
DR InterPro; IPR00504; RNA\_rec\_mot.  
DR Pfam; PF00076; rim; 2.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 250 AA; 28444 MW; ED40BD015DSFA31A CRC64;

Query Match 51.6%; Score 748; DB 10; Length 250;  
Best Local Similarity 57.6%; Pred. No. 1.5e-58;  
Matches 167; Conservative 22; Mismatches 41; Indels 60; Gaps 7;

Qy 1 MSSRNRTIYVGNLPGDIRKEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDPDADD 60  
Db 1 MSRENSRTIYVGNLPGDIREREVEDLFYKGRILIDLKIPRPPGYAF--FEDPDADD 58  
Qy 61 AIYGRDGYDFGCRRLRVEIAHGGRRFSPVDRYSSYSASR--APRRSDYRVLVTGLPP 118  
Db 59 AICRGDGYDFGCRRLRVEIAHGGRRFSPVDRYSSYSASR--APRRSDYRVLVTGLPP 117  
Qy 119 SASWQDLKDHMRKAGDYCFSEVFPDRKMGSGVDYSNYDDMKYAIRKLDATFRNAPSSA 178  
Db 118 SASWQDLKDHMRKAGDYCFSDVVEAGATGIVDYTYIEDMKYAIRKLDSEFRNAPSSA 177  
Qy 179 YIRV-----REYRSRSVSRPDDSKSVRSRSRGRGSCSVSSKRSRVSPPA 223  
Db 178 YIRIICLDAPDCDLSTCSFYDRSVSRSP-----SPV 211  
Qy 224 RSISPRPLSRPLSRSLYSVSRSGSLLRAGDWISQSRKSRSRSSNSPV 273  
Db 212 -----DERSISRSRTVPSSPSRGRSV-----SKSPRSJSRSPSPV 247

RESULT 13  
Q9SPF13 PRELIMINARY; PRT; 261 AA.  
AC Q9SPF13;  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE Splicing factor SR1D.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Landsberg;  
RA Lazar G., Goddard H.M.;  
RT "The Arabidopsis splicing factor SR1 is regulated by alternative  
splicing."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF173640; AAD52612.1; -  
DR HSSP; P19339; 1GXL.  
DR InterPro; IPR00504; RNA\_rec\_mot.  
DR Pfam; PF00076; rim; 2.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 261 AA; 28848 MW; A262A497D9BC48BE CRC64;

Query Match 51.3%; Score 745; DB 10; Length 261;  
Best Local Similarity 55.4%; Pred. No. 2.9e-58;  
Matches 150; Conservative 33; Mismatches 48; Indels 40; Gaps 4;

Qy 1 MSSRNRTIYVGNLPGDIRKEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDPDADD 60  
Db 1 MSSRSRTIYVGNLPGDIREREVEDLFYKGPVVQIDLKIPRPPGYAFVEFDARDAED 60  
Qy 61 AIYGRDGYDFGCRRLRVEIAHGGRRFSPVDRYSSYS--SRAP 103  
Db 61 AIHGRDGYDFGCRRLRVEIAHGGRR---SSDTRGSFNGGGRGGRGGRGGRGGRG 117

Qy 104 SRSYRVLVTGLPPSASWQDLKDHMRKAGDYCFSEVFPDRKMGSGVDYSNYDDMKYAI 163  
Db 118 SRSSEFRVLVTGLPPSASWQDLKDHMRKAGDYCFSEVFPDRKMGSGVDYSNYDDMKYAI 177  
Qy 164 RKLDATFRNAPSSAYIRVREYRSRSRSPDDSKSVRSRSRGRGSCSVSSKRSRVSPPA 223  
Db 178 KLDATFRNAPSSAYIRVREYRSRSRSPDDSKSVRSRSRGRGSCSVSSKRSRVSPPA 226  
Qy 224 RSISPRPLSRPLSRSLYSVSRSGSLLRAGD 254  
Db 227 QNL-----HQDLLAPARSQGHRLQEGN 248

RESULT 14  
Q8GXSO PRELIMINARY; PRT; 178 AA.  
AC Q8GXSO;  
DT 01-MAR-2003 (TREMELrel. 23, Created)  
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN AT4G02430/T14P8.21.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
RA Hayashizaki Y., Shinozaki K.;  
RT "Arabidopsis thaliana full-length cDNA."  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK118074; BAC42705.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 178 AA; 19753 MW; DA1E728F07E2B6F3 CRC64;

Query Match 44.8%; Score 649.5; DB 10; Length 178;  
Best Local Similarity 69.8%; Pred. No. 5.4e-50;  
Matches 125; Conservative 16; Mismatches 21; Indels 17; Gaps 2;

Qy 1 MSSRNRTIYVGNLPGDIRKEVEDLFYKYGPIVDIDLKIPRPPGYAFVEFDPDADD 60  
Db 1 MSSRSRTIYVGNLPGDIREREVEDLFYKGPVVQIDLKIPRPPGYAFVEFDARDAED 60  
Qy 61 AIYGRDGYDFGCRRLRVEIAHGGRRFSPVDRYSSYS--SRAPSR 106  
Db 61 AIYGRDGYDFGCRRLRVEIAHGGRRSSHDA--RGSYSGRGRGGRGGRGGRGGRG 117  
Qy 107 SRYRVLVTGLPPSASWQDLKDHMRKAGDYCFSEVFPDRKMGSGVDYSNYDDMKYAIK 165  
Db 118 SRYRVVSGLPSSASWQDLKDHMRKAGDYCFSEVFPDRKMGSGVDYSNYDDMKYAIK 176

RESULT 15  
Q81290 PRELIMINARY; PRT; 294 AA.  
AC Q81290;  
DT 01-NOV-1998 (TREMELrel. 08, Created)  
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE T14P8.21 protein (At4G02430 protein).  
GN T14P8.21 OR At4G02430.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:49:20 ; Search time 21 Seconds  
(without alignments)  
562.130 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451

Sequence: 1 MSSRWNRIRYGVNLPDGRK.....RSKSRSRSPVSPVISG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pap:\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pap:\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pap:\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pap:\*
  - 5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pap:\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 172   | 11.9        | 75     | 1     | US-07-881-075-12   |
| 2          | 172   | 11.9        | 75     | 1     | US-08-120-827-12   |
| 3          | 172   | 11.9        | 75     | 1     | US-08-478-675-12   |
| 4          | 158   | 10.9        | 437    | 2     | US-08-935-450-8    |
| 5          | 154.5 | 10.6        | 614    | 5     | PCT-US95-03236-21  |
| 6          | 132.5 | 9.1         | 579    | 4     | US-09-643-597-348  |
| 7          | 132.5 | 9.1         | 579    | 4     | US-09-542-615A-348 |
| 8          | 132.5 | 9.1         | 579    | 4     | US-09-608-421B-348 |
| 9          | 131.5 | 9.1         | 579    | 4     | US-09-643-597-176  |
| 10         | 131.5 | 9.1         | 579    | 4     | US-09-480-884A-176 |
| 11         | 131.5 | 9.1         | 579    | 4     | US-09-542-615A-176 |
| 12         | 131.5 | 9.1         | 579    | 4     | US-09-608-421B-176 |
| 13         | 130.5 | 9.0         | 255    | 4     | US-09-370-838-183  |
| 14         | 129.5 | 8.9         | 444    | 1     | US-07-881-075-3    |
| 15         | 129.5 | 8.9         | 444    | 1     | US-08-120-827-3    |
| 16         | 129.5 | 8.9         | 444    | 1     | US-08-478-675-3    |
| 17         | 127.5 | 8.8         | 177    | 4     | US-09-370-838-202  |
| 18         | 126.5 | 8.7         | 177    | 4     | US-09-370-838-208  |
| 19         | 126   | 8.7         | 380    | 1     | US-07-881-075-51   |
| 20         | 126   | 8.7         | 380    | 1     | US-08-120-827-51   |
| 21         | 126   | 8.7         | 380    | 1     | US-08-478-675-51   |
| 22         | 125.5 | 8.6         | 1850   | 4     | US-09-620-093A-5   |
| 23         | 123.5 | 8.5         | 577    | 3     | US-09-261-885-2    |
| 24         | 121   | 8.3         | 359    | 1     | US-07-881-075-2    |
| 25         | 121   | 8.3         | 359    | 1     | US-08-120-827-2    |
| 26         | 121   | 8.3         | 359    | 1     | US-08-478-675-2    |
| 27         | 120   | 8.3         | 703    | 3     | US-08-910-925-4    |

## ALIGNMENTS

### RESULT 1

US-07-881-075-12

; Sequence 12, Application US/07881075

; Patent No. 5444149

; GENERAL INFORMATION:

; APPLICANT: KEENE, JACK D.

; APPLICANT: KING, PETER H.

; APPLICANT: LEVINE, TODD

; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE

; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS

; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

; ADDRESSEE: P.C.

; STREET: 1755 Jefferson Davis Highway, Fourth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/881,075

; FILING DATE: 19920511

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Oblon, No. 5444149man F.

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 714-154-0

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)521-4500

; TELEFAX: (703)486-2347

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 75 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-07-881-075-12

Query Match 11.9%; Score 172; DB 1; Length 75;

Best Local Similarity 48.8%; Pred. No. 5.4e-10; Indels 2; Gaps 1;

Matches 36; Conservative 11; Mismatches 25;

QY 9 IYVGNLPGRKCEVEDLVKYGPIVDIKLIPRPPGYAFVEFDPDADDAIYGRDGY 68

```

Db      2  VYVGNLSSASKHEIEGAFKYGPLRNV--WVARPPGFAFVEFEDRDAEDATRALDGT 59
QY      69  DFDGCLRLRVEIAHG 82
Db      60  RCGGTRIRVEMSSG 73

RESULT 2
US-08-120-827-12
; Sequence 12, Application US/08120827
; Patent No. 5525495
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC A
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULA
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/120,827
; FILING DATE: 15-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5525495man P.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-158-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-120-827-12

Query Match 11.9%; Score 172; DB 1; Length 75;
Best Local Similarity 48.6%; Pred. No. 5.4e-10;
Matches 36; Conservative 11; Mismatches 25; Indels 2; Gaps 1

QY      9  IVGNLPGDIRKCEVDLEFYKGPVIDDLK1PPRPGYAFVEFEDRDAEDATYGRDGY 68
Db      2  VYVGNLSSASKHEIEGAFKYGPLRNV--WVARPPGFAFVEFEDRDAEDATRALDGT 59
QY      69  DFDGCLRLRVEIAHG 82
Db      60  RCGGTRIRVEMSSG 73

RESULT 3
US-08-478-675-12
; Sequence 12, Application US/08478675
; Patent No. 5773246
; GENERAL INFORMATION:
; APPLICANT: KEENE, JACK D.
; APPLICANT: KING, PETER H.

```



```

; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Query Match
Best Local Similarity 9.1%; Score 132.5; DB 4; Length 579;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db 4 LYIGNLSENAAPSDESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESWALKAE 55
QY 64 GRDGYDFDGCRLRYEIAHGRRRFPSPVDRYSSYSASRAPSRSDYRVLTGLPPSASWQ 123
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VPKRQIRKQLQIRNIPPHLOWE 95
QY 124 LDKHMRKAGDV--CFSEVFPDRKMGVVDYSDYNDKMYAIRKLDATERN-AFSSAYI 180
Db 96 VLDSLQVGVVESC-EQVNTDSETAVVNTYSSKQDQARQALDKLNGFLENFTLKVAYI 154
QY 181 RVREYERSVSRSPDDSKSVRSRSGPSCSVSKSRSVS-PARSISPRSRPLSRSL 239
Db 155 PDETAQONPLQQRGRGLGQSGSGQSGSVSKORPCDLPLRLVLP-----TQFV 207
QY 240 YSSVSRGSLLRAGDWISQSRKSRSRSRNS 271
Db 208 GAIGKEGATIR--NITKQTSKIDVHRKENA 237

RESULT 9
US-09-643-597-176
; Sequence 176, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-176

Query Match
Best Local Similarity 9.1%; Score 131.5; DB 4; Length 579;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db 4 LYIGNLSENAAPSDESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESWALKAE 55
QY 64 GRDGYDFDGCRLRYEIAHGRRRFPSPVDRYSSYSASRAPSRSDYRVLTGLPPSASWQ 123
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VPKRQIRKQLQIRNIPPHLOWE 95
QY 124 LDKHMRKAGDV--CFSEVFPDRKMGVVDYSDYNDKMYAIRKLDATERN-AFSSAYI 180

```

```

; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Query Match
Best Local Similarity 9.1%; Score 132.5; DB 4; Length 579;
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63
Db 4 LYIGNLSENAAPSDESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESWALKAE 55
QY 64 GRDGYDFDGCRLRYEIAHGRRRFPSPVDRYSSYSASRAPSRSDYRVLTGLPPSASWQ 123
Db 56 ALSG-----KIEL-HG-----KPIEVHS-----VPKRQIRKQLQIRNIPPHLOWE 95
QY 124 LDKHMRKAGDV--CFSEVFPDRKMGVVDYSDYNDKMYAIRKLDATERN-AFSSAYI 180
Db 96 VLDSLQVGVVESC-EQVNTDSETAVVNTYSSKQDQARQALDKLNGFLENFTLKVAYI 154
QY 181 RVREYERSVSRSPDDSKSVRSRSGPSCSVSKSRSVS-PARSISPRSRPLSRSL 239
Db 155 PDETAQONPLQQRGRGLGQSGSGQSGSVSKORPCDLPLRLVLP-----TQFV 207
QY 240 YSSVSRGSLLRAGDWISQSRKSRSRSRNS 271
Db 208 GAIGKEGATIR--NITKQTSKIDVHRKENA 237

RESULT 8
US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348

```



Db 96 VLDSLIVQYGVVESC-EQVNTDSETAVVNVVYSSKDQARQALDKLNGFLENFTLKVAYI 154  
QY 181 RVREYESVSRSPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239  
Db 155 PDMAAQNPLOQPRGRGLGQSGRSQSGSVSKQKPCDPLRLVLP-----TQFV 207  
QY 240 YSVSRSGSLLRAGDWISQSRKSRSRSSNS 271  
Db 208 GAIIGEGATIR--NITKQTSKIDVHRKENA 237

## RESULT 10

US-09-480-884A-176

; Sequence 176, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 176  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-480-884A-176

Query Match 9.1%; Score 131.5; DB 4; Length 579;

Best Local Similarity 23.5%; Pred. No. 0.00012;  
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKGPVIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63  
Db 4 LYIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESALKAE 55  
QY 64 GRGDYDFDGCRLRVEIAHGGRFRFSPVDYSSYSASRAPSRSDYRVLTGLPPSASWQ 123  
Db 56 ALSG-----KIEL-HG-----KPIEVEHS-----VPKQRIKQLQIRNIPPHLQWE 95  
QY 124 DLKDHMRKAGDV--CFSEVFPDRKMGSGVVDYSDYNDKMYAIRKLDATERN-AFSSAYI 180  
Db 96 VLDSLIVQYGVVESC-EQVNTDSETAVVNVVYSSKDQARQALDKLNGFLENFTLKVAYI 154  
QY 181 RVREYESVSRSPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239  
Db 155 PDMAAQNPLOQPRGRGLGQSGRSQSGSVSKQKPCDPLRLVLP-----TQFV 207  
QY 240 YSVSRSGSLLRAGDWISQSRKSRSRSSNS 271  
Db 208 GAIIGEGATIR--NITKQTSKIDVHRKENA 237

## RESULT 11

US-09-542-615A-176

; Sequence 176, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 176  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-542-615A-176

Query Match 9.1%; Score 131.5; DB 4; Length 579;

Best Local Similarity 23.5%; Pred. No. 0.00012;  
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKGPVIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63  
Db 4 LYIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESALKAE 55  
QY 64 GRGDYDFDGCRLRVEIAHGGRFRFSPVDYSSYSASRAPSRSDYRVLTGLPPSASWQ 123  
Db 56 ALSG-----KIEL-HG-----KPIEVEHS-----VPKQRIKQLQIRNIPPHLQWE 95  
QY 124 DLKDHMRKAGDV--CFSEVFPDRKMGSGVVDYSDYNDKMYAIRKLDATERN-AFSSAYI 180  
Db 96 VLDSLIVQYGVVESC-EQVNTDSETAVVNVVYSSKDQARQALDKLNGFLENFTLKVAYI 154  
QY 181 RVREYESVSRSPDDSKSYRSRSGRGPCSYSSKSRSVS-PARSISPRSRPLSRSL 239  
Db 155 PDMAAQNPLOQPRGRGLGQSGRSQSGSVSKQKPCDPLRLVLP-----TQFV 207  
QY 240 YSVSRSGSLLRAGDWISQSRKSRSRSSNS 271  
Db 208 GAIIGEGATIR--NITKQTSKIDVHRKENA 237

## RESULT 12

US-09-606-421B-176

; Sequence 176, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 176  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-606-421B-176

Query Match 9.1%; Score 131.5; DB 4; Length 579;

Best Local Similarity 23.5%; Pred. No. 0.00012;  
Matches 64; Conservative 45; Mismatches 116; Indels 47; Gaps 12;

QY 9 IYVGNLPGDIRKCEVEDLFYKGPVIVDIDLKIPRPP-----GYAFVEFEDPRDADDAIY 63  
Db 4 LYIGNLSENAAPSDLESIFK-----DAKIPVSGPFLVKTGYAFVDCPDESALKAE 55  
QY 64 GRGDYDFDGCRLRVEIAHGGRFRFSPVDYSSYSASRAPSRSDYRVLTGLPPSASWQ 123  
Db 56 ALSG-----KIEL-HG-----KPIEVEHS-----VPKQRIKQLQIRNIPPHLQWE 95

QY 124 DLKHMKAGDV--CFSEVFPDRKMGMSGVVDYNDKMYAIRKLDATFRN-AFSSAYI 180  
Db 96 VLDLLVQYGVESG--EQVNTDSETAVVNTYSSKDOARQALDKLNGFQLENFTLKVAYI 154  
QY 181 RVREYERSVSRPDDSKSYRSRSGPCSPCSYSSKRSVS--PARSISPRSRPLSRSL 239  
Db 155 PDMAAQNPLQFRGRRGIGQGGSSRQSGVSKQPCDPLRLLLVP-----TQFV 207  
QY 240 YSSVSRSGSLLRAGDWISQSRKSRSRSRNS 271  
Db 208 GAIIGKAGATIR--NITKQTSKIDVHRKENA 237

RESULT 13  
US-09-370-838-183  
; Sequence 183, Application US/09370838  
; Patent No. 544425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/370,838  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 183  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-370-838-183

Query Match 9.0%; Score 130.5; DB 4; Length 255;  
Best Local Similarity 28.5%; Pred. No. 4.5e-05;  
Matches 39; Conservative 25; Mismatches 44; Indels 29; Gaps 7;

QY 74 RLREVAHGRRRSPSYDRYSSSYASRAPSRSDYRVLTGLPPSASQMDLKDHWR-KA 132  
Db 48 RKEKNIKGGRNREP-----PTKR--YRAITWIPFVKQSKDLVKEKV 94  
QY 133 GYCVFSEVFPDRKMS---GWDYSDYNDKMYAIRKLDATFRNAPSSAYIRVREYERS 189  
Db 95 GEVTYVELLMDAEGKRGCAVVEFKVEESMKAAEVLN---KHSLSGRPLKYKE----- 145  
QY 190 VSRSPDSDSKYSRERS 206  
Db 146 ---DPDGEHARRAQXA 159

RESULT 14  
US-07-881-075-3  
; Sequence 3, Application US/07881075  
; Patent No. 5444149  
; GENERAL INFORMATION:  
; APPLICANT: KEENE, JACK D.  
; APPLICANT: KING, PETER H.  
; APPLICANT: LEVINE, TODD  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE  
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS  
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia

QY 124 DLKHMKAGDV--CFSEVFPDRKMGMSGVVDYNDKMYAIRKLDATFRN-AFSSAYI 180  
Db 96 VLDLLVQYGVESG--EQVNTDSETAVVNTYSSKDOARQALDKLNGFQLENFTLKVAYI 154  
QY 181 RVREYERSVSRPDDSKSYRSRSGPCSPCSYSSKRSVS--PARSISPRSRPLSRSL 239  
Db 155 PDMAAQNPLQFRGRRGIGQGGSSRQSGVSKQPCDPLRLLLVP-----TQFV 207  
QY 240 YSSVSRSGSLLRAGDWISQSRKSRSRSRNS 271  
Db 208 GAIIGKAGATIR--NITKQTSKIDVHRKENA 237

RESULT 15  
US-08-120-827-3  
; Sequence 3, Application US/08120827  
; Patent No. 5525495  
; GENERAL INFORMATION:  
; APPLICANT: KEENE, JACK D.  
; APPLICANT: KING, PETER H.  
; APPLICANT: LEVINE, TODD  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE  
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS  
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:

COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/881,075  
FILING DATE: 19920511  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5444149man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 714-154-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 444 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-881-075-3

Query Match 8.9%; Score 129.5; DB 1; Length 444;  
Best Local Similarity 21.5%; Pred. No. 0.00013;  
Matches 58; Conservative 47; Mismatches 12; Indels 43; Gaps 8;

QY 9 IYVGNLPQDIRKCEVDFLYKYGPIVDIDL---KIPPRPGYAFVEFEDPRDADDAIYGR 65  
Db 112 LIVNVLQPTWSODEIRSLFVSFGVESCKLIRKVTQSLGSGYGVXYVXQDAEKAINAL 171  
QY 66 DGYDFDGCRLRVEIAHGRRRSPSYDRYSSSYASRAPSRSDYRVLTGLPPSASQMDL 125  
Db 172 NGLRLQNKTIKVIAR-----PSESSEIKGA-----NLYVSGLPKNMTQSDL 212  
QY 126 KDHMRKAGDVCFSEVFPDR-----KMGSGVVDYNDY---DMKYAIRKLDATFRNAPS 176  
Db 213 ESLFSPYGIKTIIRILDCNITDEHAAGLSKGVGFIREDFQFEADRAIKELNGTTFKXSTE 272  
QY 177 SAYIRVREYERSVSRPDDSDSKYSRERSRSGP---SCSYSSKRSVSYPARSISPRSRPL 233  
Db 273 PITVRFANNPS--SNKSNMQPLAIIAIPQNTGRGAPPANAGAAAAAAAHFNAGRY 331  
QY 234 SRSRLSYSSVS-----SGSLLRAGDW 255  
Db 332 SSVISRYSPLTSLITNGMIQNTIASSGW 361

RESULT 15  
US-08-120-827-3  
; Sequence 3, Application US/08120827  
; Patent No. 5525495  
; GENERAL INFORMATION:  
; APPLICANT: KEENE, JACK D.  
; APPLICANT: KING, PETER H.  
; APPLICANT: LEVINE, TODD  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE  
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS  
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/120,827  
FILING DATE: 15-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Obolon, No. 5525495man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 714-158-0 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)413-3000  
TELEFAX: (703)413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 444 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
US-08-120-827-3

Query Match 8.9%; Score 129.5; DB 1; Length 444;  
Best Local Similarity 21.5%; Pred. No. 0.00013;  
Matches 58; Conservative 47; Mismatches 122; Indels 43; Gaps 8;  
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPIVDIDL--KIPRPPGYAFVEFEDPRDADDAIYGR 65  
Db 112 LIVNLPDTMSQDEIRSLFVGVESECKLRDKVTGOSLGGFYXVKQEDAEKAINAL 171  
QY 66 DGYDFDGCRLRVEIAHGRRFSPVDYSSYSASRAPRRSDYRVLVLTGLPPSASWQDL 125  
Db 172 NGLRLQNTIKVSIAR-----PSESIGKA-----NLVSGLPKNTQSDLL 212  
QY 126 KDHMRKAGDVCFSEVFPDR-----KMGSGVVDYSNYD---DMKYAIRKLDATFERNAPS 176  
Db 213 ESLFSPYCKIITSRLCDNITDEHAAGLSKGVGFIRDPQFEADRAIKELNGTTPKNSTE 272  
QY 177 SAYIRVREYERSVSRSPDDSKYSRSRSGRP---SCSYSSKRSVSPARSISPRSRPL 233  
Db 273 PITVKFANNPS-SNKNMSQPLAAAYIAPONTGRGRAFFANAAAGAAAAAAAIHPNAGRY 331  
QY 234 SRSRLYSVSR-----SGSLLRAGDW 255  
Db 332 SSVISRYSLTSDLTNGMIQNTIASGW 361

Search completed: January 29, 2004, 02:57:34  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 02:50:21 ; Search time 243 Seconds  
(without alignments)  
237.364 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451  
Sequence: 1 MSSRWRTIYVGNLPGDIRK.....RSKSRSRSRNSPVPVWISG 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap:\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap:\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap:\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap:\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap:\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap:\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap:\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap:\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap:\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap:\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap:\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap:\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap:\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap:\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 1451  | 100.0       | 279    | 14 | US-10-014-927-19  |
| 2          | 582.5 | 40.1        | 248    | 14 | Sequence 19, Appl |
| 3          | 517   | 35.6        | 253    | 10 | Sequence 22, Appl |
| 4          | 517   | 35.6        | 254    | 12 | Sequence 1601, Ap |
| 5          | 277   | 19.1        | 306    | 9  | Sequence 2448, Ap |
| 6          | 275.5 | 13.0        | 94     | 12 | Sequence 601, App |
| 7          | 195.5 | 13.5        | 155    | 9  | Sequence 344, App |
| 8          | 177.5 | 12.2        | 137    | 12 | Sequence 759, App |
| 9          | 158   | 10.9        | 364    | 15 | Sequence 2470, Ap |
| 10         | 156.5 | 10.8        | 112    | 12 | Sequence 472, App |
| 11         | 153   | 10.5        | 223    | 12 | Sequence 2584, Ap |
| 12         | 152   | 10.5        | 366    | 15 | Sequence 2995, Ap |
| 13         | 151.5 | 10.4        | 3664   | 15 | Sequence 228, App |
| 14         | 146.5 | 10.1        | 250    | 15 | Sequence 423, App |
| 15         | 146.5 | 10.1        | 252    | 12 | Sequence 78, Appl |

|    |       |      |     |    |                    |
|----|-------|------|-----|----|--------------------|
| 16 | 146.5 | 10.1 | 288 | 15 | US-10-197-666A-80  |
| 17 | 139.5 | 9.6  | 330 | 9  | US-09-764-848-34   |
| 18 | 139.5 | 9.6  | 330 | 12 | US-10-222-020-34   |
| 19 | 139.5 | 9.6  | 330 | 15 | US-10-116-016-34   |
| 20 | 139.5 | 9.6  | 330 | 15 | US-10-103-313-555  |
| 21 | 138.5 | 9.5  | 270 | 12 | US-10-264-049-2740 |
| 22 | 138.5 | 9.5  | 901 | 11 | US-09-920-705-3    |
| 23 | 138   | 9.5  | 556 | 15 | US-10-097-340-147  |
| 24 | 138   | 9.5  | 587 | 12 | US-10-313-986-501  |
| 25 | 138   | 9.5  | 594 | 12 | US-10-137-870-10   |
| 26 | 138   | 9.5  | 594 | 12 | US-10-140-018-10   |
| 27 | 138   | 9.5  | 594 | 12 | US-10-140-021-10   |
| 28 | 138   | 9.5  | 594 | 12 | US-10-140-274-10   |
| 29 | 138   | 9.5  | 594 | 12 | US-10-140-471-10   |
| 30 | 138   | 9.5  | 594 | 12 | US-10-140-807-10   |
| 31 | 138   | 9.5  | 594 | 12 | US-10-140-922-10   |
| 32 | 138   | 9.5  | 594 | 12 | US-10-140-924-10   |
| 33 | 138   | 9.5  | 594 | 12 | US-10-140-926-10   |
| 34 | 138   | 9.5  | 594 | 12 | US-10-141-698-10   |
| 35 | 138   | 9.5  | 594 | 12 | US-10-141-702-10   |
| 36 | 138   | 9.5  | 594 | 12 | US-10-141-704-10   |
| 37 | 138   | 9.5  | 594 | 12 | US-10-142-431-10   |
| 38 | 138   | 9.5  | 594 | 12 | US-10-142-432-10   |
| 39 | 138   | 9.5  | 594 | 12 | US-10-142-767-10   |
| 40 | 138   | 9.5  | 594 | 12 | US-10-143-033-10   |
| 41 | 138   | 9.5  | 594 | 12 | US-10-144-994-10   |
| 42 | 138   | 9.5  | 594 | 12 | US-10-145-628-10   |
| 43 | 138   | 9.5  | 594 | 12 | US-10-145-631-10   |
| 44 | 138   | 9.5  | 594 | 12 | US-10-145-633-10   |
| 45 | 138   | 9.5  | 594 | 12 | US-10-145-746-10   |

ALIGNMENTS

RESULT 1  
US-10-014-927-19  
; Sequence 19, Application US/10014927  
; Publication No. US20020115180A1  
; GENERAL INFORMATION:  
; APPLICANT: Barta, Andrea  
; APPLICANT: Lopato, Sergey  
; APPLICANT: Kalyna, Maria  
; APPLICANT: Dörner, Silke  
; TITLE OF INVENTION: Splice Factor  
; FILE REFERENCE: SONN:013US  
; CURRENT APPLICATION NUMBER: US/10/014,927  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: PCT/AT00/00100  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: A 727/99  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of the unknown organism: genome  
; OTHER INFORMATION: atsrp30  
US-10-014-927-19

Query Match 100.0%; Score 1451; DB 14; Length 279;  
Best Local Similarity 100.0%; Pred. No. 1.4e-128;  
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKYGIVDIDLKIPRPPGYAFVEFDPADDD 60  
Db 1 MSSRWRTIYVGNLPGDIRKCEVEDLFYKYGIVDIDLKIPRPPGYAFVEFDPADDD 60  
Qy 61 AIYGRGYDFDGCCLRAVEIAHGRRSPSDVYSSYSASRAPSRSDYRVLVTGJPPSA 120

Db 61 AIYGRDGYDFGCLRLVEIAHGGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 120  
QY 121 SWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180  
Db 121 SWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180  
QY 181 RVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKSRSVSPARSISPRSRPLSRGRSLY 240  
Db 181 RVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKSRSVSPARSISPRSRPLSRGRSLY 240  
QY 241 SSVSRSGSLLAGDWISQSRKSRSRGRSRSNPSVPVISG 279  
Db 241 SSVSRSGSLLAGDWISQSRKSRSRGRSRSNPSVPVISG 279

RESULT 2  
US-10-014-927-22  
; Sequence 22, Application US/10014927  
; Publication No. US20020115180A1  
; GENERAL INFORMATION:  
; APPLICANT: Barta, Andrea  
; APPLICANT: Lopato, Sergio  
; APPLICANT: Kalyna, Maria  
; APPLICANT: Dörner, Silke  
; TITLE OF INVENTION: Splice Factor  
; FILE REFERENCE: SONN:01305  
; CURRENT APPLICATION NUMBER: US/10/014, 927  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: PCT/AT00/00100  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: A. 727/99  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-014-927-22

Query Match 40.1%; Score 582.5; DB 14; Length 248;  
Best Local Similarity 53.6%; Pred. No. 9.5e-47;  
Matches 134; Conservative 28; Mismatches 47; Indels 41; Gaps 8;  
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLKIPRPPGYAFVFEFDPDADDAIYGRDGY 68  
Db 18 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLKIPRPPGYAFVFEFDPDADDAIYGRDGY 77  
QY 69 DFDGCLRLVEIAHGGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 111  
Db 78 DYDGYRLRVEFPGRSGRTGRTGGGGGGGAPRGYGP-----PSRRSENRY 123  
QY 112 LVTGLPPSASWQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATF 171  
Db 124 VVSGLPSPGSGWQDLKDHMRKAGDVCYADVVDG---TGVEFVRKEDMTYAVAKLDNTKF 180  
QY 172 R-NAPSSAYIRVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKSRSVSPARSISPR 229  
Db 181 RSHEGETAYIRKVDGPRSPSYGRSRSRS-RSRSRSRSR-----NSRSRSYPRSRGSPR 236  
QY 230 SRPL-SRSRS 238  
Db 237 YSPRHSRSRS 246

RESULT 3  
US-09-925-300-1601  
; Sequence 1601, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1601  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-300-1601

Query Match 35.6%; Score 517; DB 10; Length 253;  
Best Local Similarity 52.5%; Pred. No. 1.5e-40;  
Matches 115; Conservative 30; Mismatches 38; Indels 36; Gaps 8;  
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 63  
Db 48 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 102  
QY 64 GEDGYDFDGCCLRLVEI--AHGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 121  
Db 103 GRNGDYGCCLRLVEI--AHGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 155  
QY 122 WQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180  
Db 156 WQDLKDHMRKAGDVCYADVVDGVM---VEYLRKEDMEYALRKLDDTKFRSHEGETSYI 212  
QY 181 RVREYERSVSRSPDDSKSVRSRSGRSPGSCYSKSR 216  
Db 213 RV-----YPERSTSYGSRSGRSGRSDPYQSR 241

RESULT 4  
US-10-264-049-2448  
; Sequence 2448, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 2448  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-049-2448

Query Match 35.6%; Score 517; DB 12; Length 254;  
Best Local Similarity 52.5%; Pred. No. 1.5e-40;  
Matches 115; Conservative 30; Mismatches 38; Indels 36; Gaps 8;  
QY 9 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 63  
Db 49 IYVGNLPGDIRKCEVEDLFYKYGPVVDIDLK-----IPRPPGYAFVFEFDPDADDAIY 103  
QY 64 GEDGYDFDGCCLRLVEI--AHGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 121  
Db 104 GRNGDYGCCLRLVEI--AHGRRFSPVDRYSSYSASRAPRRSDYRVLVTGLPPSA 156  
QY 122 WQDLKDHMRKAGDVCSEVFPDRKMGSGVVDYSDYDMKYAIRKLDATFERNAPSSAYI 180  
Db 157 WQDLKDHMRKAGDVCYADVVDGVM---VEYLRKEDMEYALRKLDDTKFRSHEGETSYI 213

QY 181 RVREYRSVSRSPDPSKY---RRSRSRGPGSCYSK 216  
Db 214 RV-----YPERSTSYGYSRSGSGRGRDPSYQSR 242

## RESULT 5

US-09-925-297-601  
; Sequence 601, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 601  
; LENGTH: 306  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-297-601

Query Match 19.1%; Score 277; DB 9; Length 306;  
Best Local Similarity 38.8%; Pred. No. 8.2e-18;  
Matches 85; Conservative 27; Mismatches 65; Indels 42; Gaps 7;  
QY 93 YSSSYASRAPSRRSDYRLVTGLPFSASWDLKDHMKAGDVCSEVFPDRKMGSGVVD 152  
Db 20 YRRSGDKYGPTRTEYLLIVENLSSRCWSQDLKDYMFQAGEVTYADAHKGRKN-EGVIE 78  
QY 153 YSNYDMKYAIRKLDATFR-----NAPSSAYIRVEYSRVSRS----- 193  
Db 79 FVSYDMKRALEKLDGTEVNGRKIRLVEDKPGSRRRSYSRSH-SRSRSRSHRSKSR 137  
QY 194 --PDDSKYSRSGRPGSCYS-SKRSVSPARSISPR-----SRPLSRSLYSVS 244  
Db 138 SRSGSKSHSKRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRSRSGRS 197  
QY 245 RGSLLRAGDWI-----SQRSKRSRSGRSRS 271  
Db 198 RSKSKQAEKIQNDNVGPKRSRSPSRHKSRSRSGRS 236

## RESULT 6

US-09-864-408A-344  
; Sequence 344, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shimkets, Richard A.  
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 344  
; LENGTH: 94  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid  
US-09-864-408A-344

Query Match 19.0%; Score 275.5; DB 12; Length 94;  
Best Local Similarity 50.0%; Pred. No. 2.3e-18;  
Matches 68; Conservative 8; Mismatches 17; Indels 43; Gaps 4;  
QY 134 DVCSEVFPDRKMGSGVWDYSNYDDMKYAIRKLDATFRNAPSSAYIRVEYSRVSRS 193  
Db 2 DVCSEVYEGGTGIVDTYNNYDDMKYAIRKLDATFRNAPGRAYIRKEY----- 53  
QY 194 PDDSKYSRSGRPGSCYSKRSVSPARSISPRSLSRSLYSRSLYSRSGSLLRAG 253  
Db 54 --NGKGRYSRSGRSPRSY-SKRSRPSK-----SPRTR----- 84  
QY 254 DWISQSRKSRSGRSRS 269  
Db 85 -----RSSRSRSGRS 94

## RESULT 7

US-09-925-302-759  
; Sequence 759, Application US/09925302  
; Patent No. US20020044941A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925,302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 759  
; LENGTH: 155  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (147)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-302-759

Query Match 13.5%; Score 195.5; DB 9; Length 155;  
Best Local Similarity 37.1%; Pred. No. 1.6e-10;  
Matches 46; Conservative 18; Mismatches 33; Indels 27; Gaps 3;  
QY 9 IYVGNLPGDIRKCEVEDLFYKGYPIVDIDLKIPRPPGYAFVEFEDPRDADAIYGRDGY 68  
Db 43 VYVGNLGNNGKTELEAFGYGFLRSV--WVARNPPGFVAFVEFEDPRDADAVRELDGR 100  
QY 69 DFDGCRRLRVEIAHGRRFPSPVDYSSYSASRAPSRRSDYRLVTGLPFSASWQDLKDH 128  
Db 101 TLGCRVRLSNGEKR-----SRNRGP-----PPSWGRRPRDDY 135  
QY 129 MRKA 132  
Db 136 RRS 139

## RESULT 8

US-10-104-047-2470  
; Sequence 2470, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:



```

95 KUHGVNINVE-----ASKNKSXTST-KUHGVNISTCTNKELRK 133
129 MKRAGVCFSEVFPDRKMGSGVVDNMKYAIRKDATEPKN-----AFSAVIRV 183
134 FBEYGFVTECDIVK----YAFVHMERABDAVEAIRGLDNTFQGRMHVQJSTSLRTA 189
184 EYESRSVRSRPPDSKSYRSRSR--SRGPSCSYSKSR 218
190 PWYGR-----PWLLSVMERALVQVPSRSGYSGCR 221

```

RESULT 12

```

US-10-153-668-228
; Sequence 228, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: State Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260691
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 228
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-228

```

RESULT 13

```

RESULTS: 13
US-10-177-293-423
; Sequence 423, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei

```

```

; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoerscht, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 423
; LENGTH: 3664
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-423

Query Match 10.4%; Score 151.5; DB 15; Length 3664;
Best Local Similarity 26.6%; Pred.No. 0.00017;
Matches 85; Conservative 38; Mismatches 128; Indels 69; Gaps 13;

QY 4 RWNRTTYGVNLPGDIRKVEEDLFYKGPVIDIDIKIPRPPG-----YAFVEFEDPRDAD 59
DB 3 RETHLWVGNLENVREEKIIEHFXYGEVSV--KILPKGSEGGAFAFVDFVDIKSAQ 60
QY 60 DAIRG-----RDGVDFDCRLRVFIANGRRFSPSVDRYSSSVSASAPRRRSYR 110
DB 61 KAHNSVNMKGDRDLRTDYNEP-----TIPSAARGLDITVSIASRSREVSGFRGGGGPAY- 116
QY 111 VLVTGLPPASWQD-----LKDHNRKAGDVCFSEVFPDRKMGSGVVDYSNYDD--- 158
DB 117 ----GPPPSLHAREGRYERRLDGASDNREYAE--HSAIGHHERGTGDFDTRHYDQYY 170
QY 159 -----MKYAIRKLDATEF-----RNAFSSAYIRVREYESRSVSRSPDD 196
DB 171 ROPRERTLOHGLYIASRSPSNRFAHPDYEPRAREQFTLPSVVRHDIYRDDITREVRG 230
QY 197 SKSYRS--RSRGRGFCSYSSKSRSVSPARSISPRSLRSRSLYSVSRSGSLLRAGD 254
DB 231 RRPERNYQHSRSP---HSSQSRNQSPORLASQASRP-TRSPSGSGSRSSSS-----SD 282
QY 255 WISQSRKSRSRSRSPSVS 274
DB 283 SISSSSSSTSDSSDSSSSSS 302

RESULT 14
US-10-197-666A-78
; Sequence 78, Application US/10197666A
; Publication No. US20030092037A1

```



```

; GENERAL INFORMATION:
; APPLICANT: ASAHI KASEI KABUSHIKI KAISYA
; TITLE OF INVENTION: Elki phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1.
; SEQ ID NO 78
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-197-666A-78

```

```

Query Match      10.1%; Score 146.5; DB 15; Length 250;
Best Local Similarity 27.5%; Pred. No. 1.3e-05;
Matches 65; Conservative 26; Mismatches 66; Indels 79; Gaps 11;

QY      7 RTIVVGNLPGDIRK-----EVEDLFYKGPVIDIDL---XIPPRPGVAF 49
Db      66 RRRHVGNEAPDPNCCIGVGLSLYTTTERDLREVFYSKYGLPIADVISYVDDQSRASRGFAP 125

QY      50 VEPEDPDADDAIYGRDYPDGCRLRVEIAHGGRFPSPVDVYSS---SYSASAPARRS 107
Db      126 VYFENVDDAKEAKERANGMELDGRIIRVDSEITKRPHPTPTGIVMGRTPTGYS---SRRR 181

QY      108 DYRVLVITGLPPSASQDLKOHMRKAGDVCSEVFDRKMGMSGVVDYSNDDMKYAIKLD 167
Db      182 DY-----DRGYDRG-----YDRDYYSR--- 200

QY      168 ATEFR-----NAFSSAYIRVREYESRSVSRSPDDSK-SYRSRGRSRGSPCSYSK 216
Db      201 --SVYGGGGGGGGVAAADRDPQIVRR--SPSPYSGRGVSRSSSR-----SVSPR 248

```

```

RESULT 15
US-10-108-260A-4815
; Sequence 4815, Application US/10108260A
; Publication No. US2004000560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4815
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4815

```

```

Query Match      10.1%; Score 146.5; DB 12; Length 552;
Best Local Similarity 27.5%; Pred. No. 1.3e-05;
Matches 65; Conservative 26; Mismatches 66; Indels 79; Gaps 11;

QY      7 RTIVGNLPGDIRK-----EVEDLFYKGPVVDIDL---KIPRPQGYAF 49
      |||      |||      |||      |||      |||      |||
Db      68 RRRHVGNRANDPNCCLGVFLSLVTTERRDLEVFSGYKPIADYSIVVDOSRRSRGF 127

```

|    |     |                                                           |               |     |
|----|-----|-----------------------------------------------------------|---------------|-----|
| Qy | 50  | VFEFPRADDAIYORDGYDPDGCRLAVETAHGRRRSPSVDRYSS----           | SVSASASAPSRSS | 107 |
| Db | 128 | VYFENVDDAKEAKERANGMELDGRRIYVDSITKRPHPTFGIYMGRTYGS----     | SRRR          | 183 |
| Qy | 108 | DYRLVTGLPPSASWQDLKDMRKAGDVCVSEVPPDRKMGSGVVDYSDYDDMKYATRKL | 167           |     |
| Db | 184 | DYV-----DRGYDRG-----YDDRDIYSR-----                        | 202           |     |
| Qy | 168 | ATEFR-----NAPSSAIVREVESVSRSPPDSK-SYRSRSRSGPSCSYSSK        | 216           |     |
| Db | 203 | --SYRGGGGGGGWAAQDRDQIYRRR--SPSPYISRGYISRSRSGR--SYSPR      | 250           |     |

Search completed: January 29, 2004, 03:01:47  
Job time : 243 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 20:16:44 ; Search time 64 Seconds

(without alignments)  
1531.048 Million cell updates/sec

Title: US-10-014-927-19MOD\_COPY\_1\_222

Perfect score: 222

Sequence: 1 MRSRNRRTIYVGNLPGDIRK.....RRSRGRPCSYSSKRSRVSVP 222

Scoring table: UNITARY2

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/BAUM927/runat.04022004.131334.2726/app.query.fasta\_1.391  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=BAUM927 @CGN 1.56 @runat.04022004.131334.2726 -NCPUS=6 -ICPU=3 -NO\_WMAP  
-LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/8A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/8B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCUTS COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

*100% for detected items + 6 WMS*

| Result No. | Score | Query Match | Length  | ID | Description         |
|------------|-------|-------------|---------|----|---------------------|
| 1          | 48    | 21.6        | 284     | 4  | US-09-313-294A-5971 |
| 2          | 48    | 21.6        | 888     | 4  | US-09-313-294A-5971 |
| 3          | 48    | 21.6        | 1824    | 4  | US-09-313-294A-5971 |
| 4          | 48    | 21.6        | 4403765 | 3  | US-09-103-840A-2    |
| 5          | 48    | 21.2        | 4411529 | 3  | US-09-103-840A-2    |
| 6          | 47    | 21.2        | 955     | 4  | US-09-313-294A-5971 |
| 7          | 47    | 21.2        | 1336    | 4  | US-09-313-294A-5971 |
| 8          | 47    | 21.2        | 2508    | 4  | US-09-313-294A-5971 |
| 9          | 47    | 21.2        | 11282   | 4  | US-09-313-294A-5971 |
| 10         | 46    | 20.7        | 771     | 4  | US-09-313-294A-5971 |
| 11         | 46    | 20.7        | 888     | 4  | US-09-313-294A-5971 |
| 12         | 46    | 20.7        | 1420    | 4  | US-09-180-109A-23   |

|      |    |      |         |   |                      |                    |
|------|----|------|---------|---|----------------------|--------------------|
| c 13 | 46 | 20.7 | 2409    | 4 | US-09-230-225B-3     | Sequence 3, Appli  |
| 14   | 46 | 20.7 | 3489    | 2 | US-08-728-323A-1     | Sequence 1, Appli  |
| 15   | 46 | 20.7 | 3489    | 4 | US-09-298-568-1      | Sequence 1, Appli  |
| 16   | 46 | 20.7 | 3489    | 4 | US-09-410-399-1      | Sequence 1, Appli  |
| c 17 | 46 | 20.7 | 4897    | 6 | 5196516-7            | Patent No. 5196516 |
| 18   | 46 | 20.7 | 5207    | 2 | US-09-858-664A-1     | Sequence 1, Appli  |
| 19   | 46 | 20.7 | 6638    | 2 | US-08-070-301-2      | Sequence 2, Appli  |
| c 20 | 46 | 20.7 | 32207   | 2 | US-08-770-379-20     | Sequence 20, Appli |
| c 21 | 46 | 20.7 | 32207   | 3 | US-08-757-669A-20    | Sequence 20, Appli |
| c 22 | 46 | 20.7 | 32207   | 4 | US-09-330-371A-20    | Sequence 20, Appli |
| 23   | 46 | 20.7 | 77536   | 4 | US-09-410-551B-1     | Sequence 1, Appli  |
| 24   | 46 | 20.7 | 4403765 | 3 | US-09-103-840A-2     | Sequence 1, Appli  |
| 25   | 46 | 20.7 | 4411529 | 3 | US-09-103-840A-2     | Sequence 1, Appli  |
| c 26 | 45 | 20.3 | 840     | 4 | US-09-252-991A-13081 | Sequence 2027, Ap  |
| c 27 | 45 | 20.3 | 867     | 2 | US-08-961-858-2      | Sequence 4, Appli  |
| c 28 | 45 | 20.3 | 867     | 2 | US-08-961-858-4      | Sequence 2, Appli  |
| c 29 | 45 | 20.3 | 867     | 3 | US-09-089-593-2      | Sequence 4, Appli  |
| c 30 | 45 | 20.3 | 867     | 3 | US-09-089-593-4      | Sequence 3, Appli  |
| c 31 | 45 | 20.3 | 867     | 3 | US-08-950-925-3      | Sequence 3, Appli  |
| c 32 | 45 | 20.3 | 867     | 4 | US-09-565-286-3      | Sequence 1330, Ap  |
| 33   | 45 | 20.3 | 1065    | 4 | US-09-252-991A-13081 | Sequence 11557, A  |
| 34   | 45 | 20.3 | 1125    | 4 | US-09-252-991A-11557 | Sequence 9574, Ap  |
| c 35 | 45 | 20.3 | 1203    | 4 | US-09-252-991A-9574  | Sequence 4120, Ap  |
| c 36 | 45 | 20.3 | 1404    | 4 | US-09-252-991A-4420  | Sequence 12931, A  |
| c 37 | 45 | 20.3 | 1551    | 4 | US-09-252-991A-12931 | Sequence 16490, A  |
| c 38 | 45 | 20.3 | 1620    | 4 | US-09-252-991A-16490 | Sequence 13081, A  |
| c 39 | 45 | 20.3 | 1659    | 4 | US-09-252-991A-13081 | Sequence 188, App  |
| c 40 | 45 | 20.3 | 1845    | 4 | US-09-614-034-188    | Sequence 12858, A  |
| c 41 | 45 | 20.3 | 2073    | 4 | US-09-252-991A-12858 | Sequence 16218, A  |
| c 42 | 45 | 20.3 | 2133    | 4 | US-09-252-991A-16218 | Sequence 9636, Ap  |
| c 43 | 45 | 20.3 | 2346    | 4 | US-09-252-991A-9636  | Sequence 16381, A  |
| 44   | 45 | 20.3 | 2358    | 4 | US-09-252-991A-16381 | Sequence 4186, Ap  |
| 45   | 45 | 20.3 | 2454    | 4 | US-09-252-991A-4186  |                    |

ALIGNMENTS

RESULT 1  
US-09-313-294A-5971  
; Sequence 5971, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalguadi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 5971  
; LENGTH: 284  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700350992H1  
; NAME/KEY: unsure  
; LOCATION: 70, 83, 239, 258-259, 283  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-5971

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Alignment Scores:      | 0.000278 | Length:       | 284 |
| Pred. No.:             | 48.00    | Matches:      | 38  |
| Score:                 | 69.57%   | Conservative: | 10  |
| Best Local Similarity: | 55.07%   | Mismatches:   | 21  |
| Query Match:           | 21.62%   | Indels:       | 0   |
| DB:                    | 4        | Gaps:         | 0   |

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-313-294A-5971 (1-284)

Qy 39 LysileProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAla 58  
Db 3 AAGTCCCCCAAGACCACTGGTATGCTTTTGTGATTTGAGATTCCTCGTGATGCT 62  
Qy 59 AspAspAlaileTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGlu 78  
Db 63 GAGGAGGNAATTGCTGAACGNGATGGATACAACTTTGATGACACCGCTCTAAGAGTGGAG 122  
Qy 79 IleAlaHisGlyGlyArgArg\*\*\*\*\*SerTyrSer 98  
Db 123 GCTGCTCATGGTGTAGAGTAATGCTTCCTCGCATGATCGTTCAAGTGGCTTGGTGCG 182  
Qy 99 AlaSerArgAlaProSerArgArgSer 107  
Db 183 GTGGTGGAGCAGTACGTGGTGTGTCG 209

RESULT 2  
US-09-252-991A-12914  
; Sequence 12914, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12914  
; LENGTH: 888  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (111)  
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-12914

Alignment Scores:  
Pred. No.: 0.0887 Length: 888  
Score: 48.00 Matches: 38  
Percent Similarity: 21.92% Conservative: 10  
Best Local Similarity: 17.35% Mismatches: 171  
Query Match: 21.62% Indels: 0  
DB: 4 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-252-991A-12914 (1-888)

Qy 3 SerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22  
Db 198 AGCGGTACACGCGCGTGTGGCGTCGATCCCGACGCTCCAGCAGGAAACAGCCGAGCAGCG 257  
Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAlaPheAspLeuLysIleProPro 42  
Db 258 CCGCCAGCAGGCTTTCAGCGCGCGTCCCGGCCATCCCGCGGAGGAGCAGTGGCGA 317  
Qy 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaile 62  
Db 318 ACACCATCAGCAGCAATATTCGCGGCGGACCGAGCGATCCGACCTTGGCCAGCAGCG 377  
Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82  
Db 378 GGGCGAAGCAGGACCATCCCGCAGGTGGCGATGAAAGCGCCCAATCAACCACTCCAGGCGC 437  
Qy 83 GlyArgArg\*\*\*\*\*SerTyrSerAlaSerArgAla 102  
Db 438 ACAGCGACAGGCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 497  
Qy 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyr 122

Db 498 TCATCAGGTCGAGGCTTCGCGGGATGTTGAGCAGGATCGAGTGTATCCCGCCGCGCT 557  
Qy 123 GluAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142  
Db 558 ATTTCGAGCCCAAGGTATACGCGCGGAGCAGGATCAGCGCGGTTTCGCGCGGCGAGCGGA 617  
Qy 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla 162  
Db 618 GGGCGAATCCACCGGATGACAGTGCACCGCTTGTATCGACCGAGGCTTGGCAGCA 677  
Qy 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182  
Db 678 GGCTTACGATGTCGCGATCAGGTGCGGATCAGCGCGGTGATCAGTGTGGTGGCGGTCA 737  
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202  
Db 738 TCGCAGCGCGGAAGCGCTGCGGAGTAGTTCAAGGTTTCCATCAGCAAGTCTCCACAGC 797  
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221  
Db 798 GGCGAAGTCCAGAACGCCCGCAGCGGAGCGTCCAGGCGCGGTGCGAACAGC 854

RESULT 3  
US-09-252-991A-12570/c  
; Sequence 12570, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12570  
; LENGTH: 1824  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (976),(1124)  
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-12570

Alignment Scores:  
Pred. No.: 3.38 Length: 1824  
Score: 48.00 Matches: 38  
Percent Similarity: 21.92% Conservative: 10  
Best Local Similarity: 17.35% Mismatches: 171  
Query Match: 21.62% Indels: 0  
DB: 4 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-252-991A-12570 (1-1824)

Qy 3 SerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22  
Db 889 AGCGGTACACGCGCGTGTGGCGTCGATCCCGACGCTCCAGCAGGAAACAGCCGAGCAGCG 830  
Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAlaPheAspLeuLysIleProPro 42  
Db 829 CCGCCAGCAGGCTTTCAGCGCGCGTCCCGGCCATCCCGCGGAGGAGCAGTGGCGA 770  
Qy 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaile 62  
Db 769 ACACCATCAGCAGCAATATTCGCGGCGGACCGAGCGATCCGCCACTTGGCCAGCAGCG 710  
Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82



```
QY 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42
Db 2438026 TGCCECCGCTGGCGATCTTGGCCCTGGACCAATAGCGCTGTCCAGCAGCGCGCTCTCCA 2437967
QY 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
Db 2437966 ACGGTCCTCTGTCGACCTTCGACCGGGGTATGTAAGTGGCGCGCGCATGGAT 2437907
QY 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
Db 2437906 GCAGAACGGAAGCTTAAGCGCATGCTGGCGCGCACCGAGCTCTACACTTGGCGGGT 2437847
QY 83 GlyArgArg*****SerTyrSerAlaSerArgAla 102
Db 2437846 GCGCAGCATCTCCGCTGGCGATGACGTTTGGACCCCTCAGCAACCGACGACGACCTGCC 2437787
QY 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
Db 2437786 CCGGTCGCCGAGCTGCTCGCGCTTCCGGTCAGCAAGTGGCCAGCGCTGGCGGAAGG 2437727
QY 123 GluAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
Db 2437726 TCATCTGGTGGCGCTGGCGCGCTGGTGGCGTGGTATCCCTCAGCTCTCTTCACCAA 2437667
QY 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla 162
Db 2437666 CTCGGTCAGGTGGTCAAGAGCCTGCCGGCTGTCACCATCTGCACGACGGGGCTA 2437607
QY 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
Db 2437606 TCGCGACACCGAGATTATCGCTGGCTGTTTACCCCGCATCCGCTGCATCCATCCCG 2437547
QY 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
Db 2437546 CGACGGCTCGGTGACGGCGCTCAGCAACGCCCGCTCCGCTCGATGCATGATGCCACCA 2437487
QY 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 2437486 GGCTCGCAGGTGGTGGCGCGCGCCAGGCTATGGCGTATTGACCACCGGGT 2437433
```

RESULT 6

```
US-09-620-312D-524
; Sequence 524, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
```

```
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 524
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(866)
US-09-620-312D-524

Alignment Scores:
Pred. No.: 0.298 Length: 955
Score: 47.00 Matches: 37
Percent Similarity: 22.17% Conservative: 10
Best Local Similarity: 17.45% Mismatches: 165
Query Match: 21.17% Indels: 0
DB: 4 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-620-312D-524 (1-955)
QY 11 ValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyrLysTyr 30
Db 129 GTGCGGAGCGCCACAGGCTGAGGACTTGGCCGTGAGTTGGTCGATATGGCCCTATA 188
QY 31 GlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAlaPheVal 50
Db 189 GTAGACGTTTACATTCCACTTTCACCTCGCCGCCCAAGAGGATTTGCTTATGTT 248
QY 51 GluPheGluAspProArgAspAlaIleTyrGlyArgAspGlyTyrAspPhe 70
Db 249 CAATTGAGATGTTTCGAGATGCTGAAGATGCTCTTTATTAACCTCAATAGAAAGTGGTA 308
QY 71 AspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 90
Db 309 TGTGCGCTCAGATTGAAATACAGTTTCACAGGTGATCGCAAAACACCAGGCCAAATG 368
QY 91 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
Db 369 AAATCAAAAGAACGTCATCCTTGTCTCAAGTGTACAGGAGATCAAGAGCCCGCAGC 428
QY 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130
Db 429 CAAAGAAAGACTCGAAGTAGAAGTTCTTCATGGGGAAGAAATAGGAGCGGTTCAGACAGC 488
QY 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
Db 489 CTTAAAGAGTCTCGACACAGGGGATTTCTTATAGCCAGTCTAAATCTCGTCCAAATCA 548
QY 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170
Db 549 TTACCAAGCGCGTCTACCTCAGCAAGGCGAGTCAAGAACTCCAAAGAGGAATTTTGGCTCT 608
QY 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
Db 609 AGAGGACGGTCAAGGTCCTCAAGTCTTACAAAAGAGGTCCAAAGTCAATAGGAAATACAC 668
QY 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210
Db 669 TCAAGTTCACCTCAAAAGCAGACTAGCTCAGGAACAAATCAAGATCATCATGGAAGACAT 728
QY 211 CysSerTyrSerSerLysSerArgSerValSerPro 222
Db 729 TCTGACTCAATAGCAAGATCCCGGTGTAATCTCC 764
```

RESULT 7

```
US-09-180-109A-28
; Sequence 28, Application US/09180109A
; Patent No. 6410293
; GENERAL INFORMATION:
; APPLICANT: MUKUMOTO, Fujio
; APPLICANT: NISHIO, Shoichi
; APPLICANT: AKIMARU, Jiro
; APPLICANT: MITSUDA, Satoshi
```

RESULT 8  
US-09-252-991A-793/c  
; Sequence 793, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 793  
; LENGTH: 2508  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-793

Alignment Scores:  
Pred. No.: 39.4 Length: 2508  
Score: 47.00 Matches: 37  
Percent Similarity: 21.76% Conservative: 10  
Best Local Similarity: 17.13% Mismatches: 169  
Query Match: 21.17% Indels: 0  
DB: Gaps: 4

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-180-109A-28 (1-1336)

QY 6 AsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAsp 25  
DB 6 AACAGGACTATCAGGCACTTACGATGCCGGGTACAGCGATTTTCGGTCCCGCACCA 65  
QY 26 LeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArgPro 45  
DB 66 ATCTTGTAAGCGCGGAGGATGCTAAGCTGCTGGGACATATATGCGCCCGAGG 125  
QY 46 GlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArg 65  
DB 126 CGGGCGAANTCAGAGGACGACACGTGATGACGACGACACCGCGCTGAGTCCGAGCGAC 185  
QY 66 AspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg 85  
DB 186 CCGCGCACCGACTGACCGCGCGGAGATCGCGGCTGTCGACCTGCCCTTCCACGA 245  
QY 86 \*\*\*\*\*SeryTyrSerAlaSerArgAlaProSerArg 105  
DB 246 GCTGTGTTCCGCGCGCGGATGTCACCGCGCATCAGCGCGGATCAGGTTACGCT 305  
QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125  
DB 306 GTCCAGCTGTTGTCATCAAGACGGCGGCTGCCCGAGGATTGGGCTATTGACGCA 365  
QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145  
DB 366 GTCCAGCCATCGCATACCGGCTGAAGCGACCAAGCTGATGAGCCCGCGCGTCT 425  
QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165  
DB 426 GCAGGCGGCGCGCAGCCCAAGGATCAGCGCTCGACGCGCTTCTGTCATGGGCGCGCTG 485  
QY 166 LeuAspAlaThrCluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185  
DB 486 GGGCAACCCCAAGGATCGCGACATGCCCATCGTGGAGATGGTGAAGGGCGTCCGCGC 545  
QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205  
DB 546 CATGGGCATGGAACCTGATGACGCTGGGATGTCGACCGATGACAGGCGCACACGCT 605  
QY 206 SerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220  
DB 606 CGCCGAGGCGGCGCTGACTATTACAAATCAATATCGACACGTC 650

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-252-991A-793 (1-2508)

QY 7 ArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26  
DB 1322 CGCAGGCACTCGGTTCTCGAAGCTGAGCGGCCCTCTCGCTATTGCGCGCACGACGCAA 1263  
QY 27 PheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly 46  
DB 1362 AGATTCAACAGCGGTCATGTCAGCGCGGCTCGAAGTCGGCGACGCGAGGTCCACG 1203  
QY 47 TyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAsp 66  
DB 1202 AGTTGGCGCTGGCATGCGCGAGCATCTGTTGCGCGCGGCTTGGCGTAGCGTACGCGG 1143  
QY 67 GlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg\*\*\* 86  
DB 1142 CTGTCCAGTTGACCCAGAGGATGCCAGGGTGTGTTGTCGATGGAGAACTGGGTCCAGG 1083  
QY 87 \*\*\*\*\*SeryTyrSerAlaSerArgAlaProSerArg 106  
DB 1082 CGCAGCGCTCTCGCGCGGCTGCGCGAGTCCAGGTGCGCGGCGCGCGCTACAGCGCG 1023  
QY 107 SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126  
DB 1022 GTTTCAGCGCATGCTGCTGCGCGGCTGGAGAAACACCGCAGCGCGCGCGCGCGCAGC 963  
QY 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146  
DB 962 AGCAGGCGCGAGGAGTATCGACAGTTCTGCCAGAACCCCGCGCACTCGCGAGCGCGCGA 903  
QY 147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeu 166  
DB 902 TACTTGGTTGCAACACGCGCTCGTCAACTGGTCCAGGACTTGGCGCGACGAGGCGCGC 843  
QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186  
DB 842 AACGCTGGTTCGACCACTCGCAGTTCCGGCAGTCCGCGACGGGTGCGCACCCCGTAGC 783  
QY 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer 206  
DB 782 AGCTCGGGTAGCGGATATCGCAACACCGCTAGCTCGCGCAACTCGGTTCGCGGCTC 723  
QY 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerPro 222



Db 203 ACCAGCACACGTCGCGCTCCACGGCGGCATCCAGCAGCGCTGGTGGCAGTG 144  
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202  
Db 143 ACCCGCTGACGATCGCGGCACCTCGCGCGGCCTTCGACCTGCGAGCGCTTCGGGCAG 84  
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerLysSerArgSerVal 220  
Db 83 TAGTCCTGGATGCGCGGCATCCAGATAACGGTCGGCCTCTTCGACAGGGGTG 30  
RESULT 11  
US-09-252-991A-8235  
; Sequence 8235, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 8235  
; LENGTH: 888  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-8235  
Alignment Scores:  
Pred. No.: 0.481 Length: 888  
Score: 46.00 Matches: 36  
Percent Similarity: 21.10% Conservatives: 10  
Best Local Similarity: 16.51% Mismatches: 172  
Query Match: 20.72% Indels: 0  
Gaps: 4  
DB:  
US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-252-991A-8235 (1-888)  
Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22  
Db 176 TCGGTGCGATGCTCGCGCGCGATGAGCTGCTCTCCGCGCGCTGTCAGC 235  
Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42  
Db 236 GTCTGCTCGAAACCTCGCGCGCTCAGGTAGGCATCGACGCCGCGCATCGCTGCTG 295  
Qy 43 ArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62  
Db 296 ATATAGCCTTGGCGCGCGGTGACCGAGCGATGCGCGGATCGCTGTCACCGCATCG 355  
Qy 63 TyrArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAlaHisGly 82  
Db 356 ACCAGCAGCGGTTCACGCCCGCAGCATCGCGCATCGCGCGGAAATCGCGAGCTGC 415  
Qy 83 GlyArgArg\*\*\*\*\*SerTyrSerAlaSerArgAla 102  
Db 416 ATCGGTTCGCAAGCAACCGAGGAGACGATGGAACGGGGTTGCCGCGCTCCAAAGCGC 475  
Qy 103 ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122  
Db 476 CCTTCCACTTCGAAGCGAGAGCTCGCGGCGAGTGCACGCTTGTGTCACCTCCGATGC 535  
Qy 123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142  
Db 536 AGATCAGCGCGAGTGGTGAAGCGAGCAGACTGATATCGTTGTCAGAGGCTCTTCAGG 595  
Qy 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla 162

Db 596 CGCGTGTGCTTCATCGCGCACCGCAGGGGTCTCCACCTTCAGAAAGTAACCGTGATGC 655  
Qy 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182  
Db 656 ACCAGCACACGTCGCGCTCCAGCGCGCATCCAGCAGCGCTGGTGGCAGTG 715  
Qy 183 ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202  
Db 716 ACCCGCTGAGGATCCGCGGCACCTCGCGCGGCCTTCGACCTGCGAGCGCTTCGGGCAG 775  
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerLysSerArgSerVal 220  
Db 776 TAGTCCTGGATGCGCGGCATCCAGATAACGGTCGGCCTCTTCGACAGGGGTG 829  
RESULT 12  
US-09-180-109A-23  
; Sequence 23, Application US/09180109A  
; Patent No. 6410293  
; GENERAL INFORMATION:  
; APPLICANT: MUKUMOTO, Fujio  
; APPLICANT: NISHIO, Shoichi  
; APPLICANT: AKIMARU, Jiro  
; APPLICANT: MITSUDA, Satoshi  
; TITLE OF INVENTION: DNA Fragments Containing Biotin Biosynthetase Gene and  
; TITLE OF INVENTION: Use of the Same  
; FILE REFERENCE: 0152-0490P  
; CURRENT APPLICATION NUMBER: US/09/180,109A  
; CURRENT FILING DATE: 1998-12-03  
; PRIOR APPLICATION NUMBER: 09/047838 JAPAN  
; PRIOR FILING DATE: 1997-03-03  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: Patent in ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 1420  
; TYPE: DNA  
; ORGANISM: Sphingomonas paucimobilis  
; FEATURE:  
; OTHER INFORMATION: Strain = JCM7511  
; NAME/KEY: CDS  
; LOCATION: (223)..(1245)  
US-09-180-109A-23  
Alignment Scores:  
Pred. No.: 5.17 Length: 1420  
Score: 46.00 Matches: 36  
Percent Similarity: 21.50% Conservatives: 10  
Best Local Similarity: 16.82% Mismatches: 168  
Query Match: 20.72% Indels: 0  
Gaps: 4  
DB:  
US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-180-109A-23 (1-1420)  
Qy 7 ArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26  
Db 81 AGGACTATCAGGCATCTACGATCGCGGGTACAGCGGATTTTCGGTCCCGGCACCAATC 140  
Qy 27 PheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly 46  
Db 141 TTGTAAAGCGCGCAGGATGCTGCTGAGGTGCTGGACATATATATCCGCCCGGAGCGG 200  
Qy 47 TyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAsp 66  
Db 201 GCGAATCAGCAGGACACACGCTGATGACGACGACCGCGCTGAGCTCCGAGCGACCCC 260  
Qy 67 GlyTyrAspPheAspGlyCysArgLeuValGluIleAlaHisGlyArgArg\*\*\* 86  
Db 261 GCGCAGCAGCTGGAGCGCGCGAGATCGCGCGCTGTTTCGACCTGCCCTTCCAGGCT 320  
Qy 87 \*\*\*\*\*SerTyrSerAlaSerArgAlaProSerArgArg 106  
Db 321 GTTCTTCGCGCGCGAGGTGACCGCGCGCATCAGCCGCGCATCAGGTTTTCAGCTGTC 380  
Qy 107 SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126



Db 381 GACGCTGTTCATCAAGAGCGGGGGTGCCTCCCGAGGATTGCGGTATTGCACCCAGTC 440  
QY 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146  
Db 441 GACCCATGCCATACCGGCTCAAGCGACCAAGTGTATGACCGCGCGGTGCTGCA 500  
QY 147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeu 166  
Db 501 GCGCGCGCGCAGGCGCAAGATCAAGGCTCAGCGGCTTCGCATGGCGCGCGCTCGCG 560  
QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186  
Db 561 CAACCCAGAGTCCGATCGCCCGCATCGTGGAGATGCTGAAGGCGTCCCGCCAT 620  
QY 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSer 206  
Db 621 GGGCATGAAACCTGCATGACGCTGGCATGCTGACCGATGCACAGGCGCAGCGCTCGC 680  
QY 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220  
Db 681 CGAGCGGGGTGGACTATTACATCAATCAATATCGACACGTC 722

RESULT 13  
US-09-230-225B-3/c  
; Sequence 3, Application US/09230225B  
; Patent No. 6403362  
; GENERAL INFORMATION:  
; APPLICANT: Meiji Seika Kaisha, Ltd.  
; APPLICANT: Moriya, Tatsuki  
; TITLE OF INVENTION: Systems for the Mass Production of Proteins or Peptides by Micro  
; TITLE OF INVENTION: of the Genus Humicola  
; FILE REFERENCE: VX990054  
; CURRENT APPLICATION NUMBER: US/09/230,225B  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 3  
; LENGTH: 2409  
; TYPE: DNA  
; ORGANISM: Humicola insolens  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (389)..(457)  
; OTHER INFORMATION:  
; NAME/KEY: mat\_peptide  
; LOCATION: (458)..( )  
; OTHER INFORMATION:  
; NAME/KEY: CDS  
; LOCATION: (458)..(477)  
; OTHER INFORMATION:  
; NAME/KEY: Intron  
; LOCATION: (478)..(535)  
; OTHER INFORMATION:  
; NAME/KEY: CDS  
; LOCATION: (536)..(1029)  
; OTHER INFORMATION:  
; NAME/KEY: Intron  
; LOCATION: (1030)..(1141)  
; OTHER INFORMATION:  
; NAME/KEY: CDS  
; LOCATION: (1142)..(1761)  
; OTHER INFORMATION:  
; NAME/KEY: Intron  
; LOCATION: (1762)..(1815)  
; OTHER INFORMATION:  
; NAME/KEY: CDS  
; LOCATION: (1816)..(1989)  
; OTHER INFORMATION:  
; NAME/KEY: Intron  
; LOCATION: (1990)..(2044)  
; OTHER INFORMATION:  
; NAME/KEY: CDS

; LOCATION: (2045)..(2095)  
; OTHER INFORMATION:  
; NAME/KEY: misc\_feature  
; LOCATION: (688)..(693)  
; OTHER INFORMATION: Cleavage site SmaI  
; NAME/KEY: misc\_feature  
; LOCATION: (1253)..(1259)  
; OTHER INFORMATION: Cleavage site BamHI  
; NAME/KEY: misc\_feature  
; LOCATION: (1505)..(1510)  
; OTHER INFORMATION: Cleavage site BglII  
; NAME/KEY: misc\_feature  
; LOCATION: (1643)..(1648)  
; OTHER INFORMATION: Cleavage siteStuI  
US-09-230-225B-3  
Alignment Scores:  
Pred. No.: 74.9 Length: 2409  
Score: 46.00 Matches: 36  
Percent Similarity: 21.80% Conservative: 10  
Best Local Similarity: 17.06% Mismatches: 165  
Query Match: 20.72% Indels: 0  
DB: 4 Gaps: 0  
US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-230-225B-3 (1-2409)  
QY 11 ValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyrLysTyr 30  
Db 1230 GTTGGCGCGTGTGTTGGGATCGCCACTCGCGCTTCGAAGCGCGGCGAGCGAGTCGCG 1171  
QY 31 GlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAlaPheVal 50  
Db 1170 GTCAGGAAGTCGTAACACGACGATCGGCTGGATTGGCGTTAATTTGCCGTCCAGAAATG 1111  
QY 51 GluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe 70  
Db 1110 ATGTGGAGATGTGTGGTATGTTGATGATGTGGTGTGTCAGGGGTACAGGTTAGGTGG 1051  
QY 71 AspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg\*\*\*\*\* 90  
Db 1050 TGGTGTGACACCGCACITACCGGCATACGGGGGTTCGGCGCCCGCTGGTTCGGCGCGC 991  
QY 91 \*\*\*\*\*SerTyrSerAlaSerArgAlaProSerArgSerArgSerArg 110  
Db 990 GGATCTCAGAGAGGTCTCGACGAGCGGTGTCCACCGTGACGTTGGGTGCGAGCCACT 931  
QY 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130  
Db 930 GGAAGCTCGGACCTCAGCGACGCGCGCGCGCCCTCAAGCGAGGTCGGTGTGATCT 871  
QY 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150  
Db 870 GAGGAATGCGAGGGTGTGGACCTCAGAGCGGTAGTAGTTGTTGCCCGAGAGCTGGACAC 811  
QY 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170  
Db 810 CCTCGAAGGGTTCGCGTTGTAGTGGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 751  
QY 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190  
Db 750 TCGGGGTGGGAGCAGTGTGTGATCGAGGTCAACACCACCGGTCTGTGTGTGTGTGTGTGG 691  
QY 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgSerArg 210  
Db 690 GGGAGGTGTGTGACGAGCTCGAAGTGTGACGTAGTGTGAGGTGTGTGTGTGTGTGTGTGT 631  
QY 211 CysSerTyrSerSerLysSerArgSerValSer 221  
Db 630 GCAAGCAGTGGGAGTACCGAGTCTGTTCTGTCTCA 598  
RESULT 14  
US-08-728-323A-1  
; Sequence 1, Application US/08728323A



```

Db 1208 AGGAGGATGACGATGATGAGGACAAATGAGGACGAGGAGGATGACGAGGAGGAGGACAAAGA 1267
      |||
Qy 84 ArgArg*****SerTyrSerAlaSerArgAlaPro 103
      |||
Db 1268 AGGAGGACGAGGAGGACGGGGCGGATGGAACAAACGTTGACATCCAAAGTTCACAAC 1327
      |||
Qy 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123
      |||
Db 1328 AGCAGCAGGAGCCCAACACAGCAGGAGCCACAGCAGCAGGAGGCCACACAGCAGGAGGCCCC 1387
      |||
Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143
      |||
Db 1388 TGCAGGAGCCCAACACAGCAGGAGCCACAGCAGGAGGCCACACAGCAGGAGGCCCTGC 1447
      |||
Qy 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle 163
      |||
Db 1448 AGCAGCCACAACACAGCAGGAGCCACAGCAGCAGGAGGCCCTGCAGGAGGCCACACACAGCAGG 1507
      |||
Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183
      |||
Db 1508 AGCCACAACAGCAGGAGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACACAGCAGG 1567
      |||
Qy 184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg 203
      |||
Db 1568 AGCCACAGCAGCAGGAGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACACAGCAGG 1627
      |||
Qy 204 SerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221
      |||
Db 1628 AGCCACAGCAGCAGGAGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGGCCACAGCAGG 1681
      |||

```

Search completed: February 4, 2004, 21:53:57  
 Job time : 1275 secs

Score and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID        | Description        |
|------------|-------|-------|--------|----|-----------|--------------------|
| 1          | 222   | 100.0 | 771    | 21 | AAC49242  | Arabidopsis thalia |
| 2          | 222   | 100.0 | 925    | 21 | AAC49248  | Arabidopsis thalia |
| 3          | 222   | 100.0 | 1213   | 21 | AAC40180  | Arabidopsis thalia |
| 4          | 218   | 98.2  | 714    | 21 | AAC42865  | Arabidopsis thalia |
| 5          | 207   | 93.2  | 762    | 24 | AB213394  | Arabidopsis thalia |
| 6          | 143   | 64.4  | 1232   | 21 | AAC34814  | Arabidopsis thalia |
| 7          | 141   | 63.5  | 1190   | 21 | AAC38528  | Arabidopsis thalia |
| 8          | 130.5 | 58.8  | 1241   | 21 | AAC37026  | Arabidopsis thalia |
| 9          | 116.5 | 52.5  | 885    | 21 | AAC42654  | Arabidopsis thalia |
| 10         | 98    | 44.1  | 479    | 21 | AAC35763  | Zea mays DNA fragm |
| 11         | 95    | 42.8  | 1428   | 22 | AAC402904 | Human shear stress |
| 12         | 90    | 40.5  | 1438   | 21 | AAC44527  | Zea mays DNA fragm |
| 13         | 87.5  | 39.4  | 1069   | 24 | ABN96845  | Human prostate can |
| 14         | 87.5  | 39.4  | 1162   | 21 | AAF16226  | Human prostate can |
| 15         | 87.5  | 39.4  | 1218   | 23 | ABV30279  | Human prostate exp |
| 16         | 87.5  | 39.4  | 1475   | 24 | ABO54393  | Human ovarian anti |
| 17         | 87.5  | 39.4  | 1584   | 24 | ABY93368  | Mouse ischaemic co |
| 18         | 81    | 36.5  | 345    | 24 | ABQ85913  | Arabidopsis thalia |
| 19         | 73.5  | 33.1  | 446    | 24 | ABY94679  | Human pancreatic c |
| 20         | 73.5  | 33.1  | 1478   | 23 | ABL29431  | Drosophila melanog |
| 21         | 73    | 32.9  | 492    | 21 | AAC47771  | Zea mays DNA fragm |
| 22         | 69    | 31.1  | 759    | 24 | AB577425  | Frog embryonic gen |
| 23         | 69    | 31.1  | 2601   | 24 | AB235222  | Human gene express |
| 24         | 65    | 29.3  | 742    | 24 | AB577306  | Frog embryonic gen |
| 25         | 63    | 28.4  | 495    | 24 | ABY99907  | Mouse ischaemic co |
| 26         | 63    | 28.4  | 497    | 21 | AAC09125  | Human secreted pro |
| 27         | 62    | 27.9  | 2076   | 24 | ABK83839  | Human cDNA differe |
| 28         | 61.5  | 27.7  | 444    | 23 | ABV45807  | Human prostate exp |
| 29         | 60.5  | 27.3  | 727    | 23 | ABV16009  | Human prostate exp |
| 30         | 56    | 25.2  | 4044   | 21 | AAC81899  | A. thaliana SRP30  |
| 31         | 55.5  | 25.0  | 1454   | 24 | ABK84288  | Human cDNA differe |
| 32         | 55.5  | 25.0  | 1534   | 25 | ABX63268  | Human cDNA #268 di |
| 33         | 55.5  | 25.0  | 3697   | 25 | ABX62342  | Human activated T  |
| 34         | 55.5  | 25.0  | 3948   | 23 | ABL29430  | Drosophila melanog |
| 35         | 55    | 24.8  | 1644   | 23 | ABL30487  | Drosophila melanog |
| 36         | 55    | 24.8  | 2367   | 20 | AAZ41956  | Human myometrium t |
| 37         | 54    | 24.3  | 471    | 21 | AAC41641  | Arabidopsis thalia |
| 38         | 54    | 24.3  | 1150   | 21 | AAC51497  | Arabidopsis thalia |
| 39         | 54    | 24.3  | 1153   | 21 | AAC33304  | Arabidopsis thalia |
| 40         | 53    | 23.9  | 568    | 21 | AAC34694  | Arabidopsis thalia |
| 41         | 53    | 23.9  | 913    | 20 | AAC23485  | Human prostate can |
| 42         | 53    | 23.9  | 57248  | 24 | ABK83563  | Human cDNA differe |
| 43         | 52    | 23.4  | 913    | 21 | AAC34439  | Arabidopsis thalia |
| 44         | 52    | 23.4  | 1365   | 21 | AAC51484  | Arabidopsis thalia |
| 45         | 52    | 23.4  | 1367   | 21 | AAC33948  | Arabidopsis thalia |

ALIGNMENTS

RESULT 1  
AAC49242  
ID AAC49242 standard; DNA; 771 BP.  
XX  
AC AAC49242;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60449.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EF1033405-A2.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 20:08:19 ; Search time 266 Seconds  
(without alignments)  
2252.914 Million cell updates/sec

Title: US-10-014-927-19MOD\_COPY\_1\_222  
Perfect score: 222  
Sequence: 1 MSSRNRTIYGNLPGDIRK.....RSRGRGPCSVSSKRSRVSP 222

Scoring table: UNITARY2  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/BAUM927/runat\_04022004\_131333\_2700/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_19Jun03 -QFWT=fastap -SUBFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=BAUM927 @CGN 1.1.312 @runat\_04022004\_131333\_2700 -NCFU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-embl/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-embl/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

|    |                              |              |               |
|----|------------------------------|--------------|---------------|
| XX | 06-SEP-2000.                 |              |               |
| PD |                              |              |               |
| XX |                              |              |               |
| XX | 25-FEB-2000; 2000EP-0301439. |              |               |
| XX |                              |              |               |
| XX | 25-FEB-1999; 99US-0121825.   | 02-JUL-1999; | 99US-0142055. |
| PR | 05-MAR-1999; 99US-0123180.   | PR           | 99US-0142390. |
| PR | 09-MAR-1999; 99US-0123548.   | PR           | 99US-0142803. |
| PR | 23-MAR-1999; 99US-0125788.   | PR           | 99US-0142920. |
| PR | 25-MAR-1999; 99US-0126264.   | PR           | 99US-0142977. |
| PR | 29-MAR-1999; 99US-0126785.   | PR           | 99US-0143542. |
| PR | 01-APR-1999; 99US-0127482.   | PR           | 99US-0143542. |
| PR | 06-APR-1999; 99US-0128234.   | PR           | 99US-0143624. |
| PR | 08-APR-1999; 99US-0128714.   | PR           | 99US-0144005. |
| PR | 16-APR-1999; 99US-0129845.   | PR           | 99US-0144086. |
| PR | 19-APR-1999; 99US-0130077.   | PR           | 99US-0144086. |
| PR | 21-APR-1999; 99US-0130449.   | PR           | 99US-0144325. |
| PR | 23-APR-1999; 99US-0130510.   | PR           | 99US-0144331. |
| PR | 23-APR-1999; 99US-0130891.   | PR           | 99US-0144332. |
| PR | 28-APR-1999; 99US-0131449.   | PR           | 99US-0144332. |
| PR | 30-APR-1999; 99US-0132048.   | PR           | 99US-0144334. |
| PR | 30-APR-1999; 99US-0132407.   | PR           | 99US-0144335. |
| PR | 04-MAY-1999; 99US-0132484.   | PR           | 99US-0144352. |
| PR | 05-MAY-1999; 99US-0132485.   | PR           | 99US-0144352. |
| PR | 06-MAY-1999; 99US-0132486.   | PR           | 99US-0144632. |
| PR | 07-MAY-1999; 99US-0132487.   | PR           | 99US-0144884. |
| PR | 11-MAY-1999; 99US-0132863.   | PR           | 99US-0144814. |
| PR | 14-MAY-1999; 99US-0134256.   | PR           | 99US-0145086. |
| PR | 14-MAY-1999; 99US-0134218.   | PR           | 99US-0145088. |
| PR | 14-MAY-1999; 99US-0134219.   | PR           | 99US-0145085. |
| PR | 14-MAY-1999; 99US-0134221.   | PR           | 99US-0145087. |
| PR | 14-MAY-1999; 99US-0134370.   | PR           | 99US-0145089. |
| PR | 18-MAY-1999; 99US-0134768.   | PR           | 99US-0145192. |
| PR | 19-MAY-1999; 99US-0134941.   | PR           | 99US-0145145. |
| PR | 20-MAY-1999; 99US-0135124.   | PR           | 99US-0145218. |
| PR | 21-MAY-1999; 99US-0135353.   | PR           | 99US-0145224. |
| PR | 24-MAY-1999; 99US-0135629.   | PR           | 99US-0145276. |
| PR | 25-MAY-1999; 99US-0136021.   | PR           | 99US-0145913. |
| PR | 27-MAY-1999; 99US-0136392.   | PR           | 99US-0145918. |
| PR | 28-MAY-1999; 99US-0136782.   | PR           | 99US-0145919. |
| PR | 01-JUN-1999; 99US-0137222.   | PR           | 99US-0145951. |
| PR | 03-JUN-1999; 99US-0137528.   | PR           | 99US-0146386. |
| PR | 04-JUN-1999; 99US-0137502.   | PR           | 99US-0146388. |
| PR | 07-JUN-1999; 99US-0137724.   | PR           | 99US-0147038. |
| PR | 08-JUN-1999; 99US-0138094.   | PR           | 99US-0147204. |
| PR | 10-JUN-1999; 99US-0138540.   | PR           | 99US-0147302. |
| PR | 10-JUN-1999; 99US-0138847.   | PR           | 99US-0147192. |
| PR | 14-JUN-1999; 99US-0139119.   | PR           | 99US-0147260. |
| PR | 16-JUN-1999; 99US-0139452.   | PR           | 99US-0147303. |
| PR | 16-JUN-1999; 99US-0139453.   | PR           | 99US-0147416. |
| PR | 17-JUN-1999; 99US-0139432.   | PR           | 99US-0147493. |
| PR | 18-JUN-1999; 99US-0139454.   | PR           | 99US-0147935. |
| PR | 18-JUN-1999; 99US-0139455.   | PR           | 99US-0148171. |
| PR | 18-JUN-1999; 99US-0139456.   | PR           | 99US-0148341. |
| PR | 18-JUN-1999; 99US-0139457.   | PR           | 99US-0148565. |
| PR | 18-JUN-1999; 99US-0139458.   | PR           | 99US-0148684. |
| PR | 18-JUN-1999; 99US-0139459.   | PR           | 99US-0149175. |
| PR | 18-JUN-1999; 99US-0139460.   | PR           | 99US-0149723. |
| PR | 18-JUN-1999; 99US-0139461.   | PR           | 99US-0149723. |
| PR | 18-JUN-1999; 99US-0139462.   | PR           | 99US-0149902. |
| PR | 18-JUN-1999; 99US-0139463.   | PR           | 99US-0149930. |
| PR | 18-JUN-1999; 99US-0139750.   | PR           | 99US-0150566. |
| PR | 21-JUN-1999; 99US-0139817.   | PR           | 99US-0150884. |
| PR | 22-JUN-1999; 99US-0139899.   | PR           | 99US-0151065. |
| PR | 23-JUN-1999; 99US-0140353.   | PR           | 99US-0151086. |
| PR | 23-JUN-1999; 99US-0140354.   | PR           | 99US-0151080. |
| PR | 24-JUN-1999; 99US-0140695.   | PR           | 99US-0151438. |
| PR | 28-JUN-1999; 99US-0140823.   | PR           | 99US-0151930. |
| PR | 29-JUN-1999; 99US-0140991.   | PR           | 99US-0152363. |
| PR | 30-JUN-1999; 99US-0141287.   | PR           | 99US-0153070. |
| PR | 01-JUL-1999; 99US-0141842.   | PR           | 99US-0153758. |
| PR | 01-JUL-1999; 99US-0141842.   | PR           | 99US-0154018. |
| PR | 01-JUL-1999; 99US-0142154.   | PR           | 99US-0154039. |
| PR |                              | PR           | 99US-0154779. |

```
PR 22-SEP-1999; 99US-01551139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 25-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 27-SEP-1999; 99US-0156596.
PR 28-SEP-1999; 99US-0157117.
PR 29-SEP-1999; 99US-0157753.
PR 30-SEP-1999; 99US-0157753.
PR 01-OCT-1999; 99US-0158029.
PR 02-OCT-1999; 99US-0158232.
PR 03-OCT-1999; 99US-0158369.
PR 04-OCT-1999; 99US-0159294.
PR 05-OCT-1999; 99US-0159294.
PR 06-OCT-1999; 99US-0159295.
PR 07-OCT-1999; 99US-0159329.
PR 08-OCT-1999; 99US-0159330.
PR 09-OCT-1999; 99US-0159331.
PR 10-OCT-1999; 99US-0159637.
PR 11-OCT-1999; 99US-0159638.
PR 12-OCT-1999; 99US-0159684.
PR 13-OCT-1999; 99US-0160741.
PR 14-OCT-1999; 99US-0160767.
PR 15-OCT-1999; 99US-0160768.
PR 16-OCT-1999; 99US-0160770.
PR 17-OCT-1999; 99US-0160814.
PR 18-OCT-1999; 99US-0160815.
PR 19-OCT-1999; 99US-0160815.
PR 20-OCT-1999; 99US-0160980.
PR 21-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 24-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 27-OCT-1999; 99US-0161359.
PR 28-OCT-1999; 99US-0161360.
PR 29-OCT-1999; 99US-0161361.
PR 30-OCT-1999; 99US-0161920.
PR 31-OCT-1999; 99US-0161920.
PR 01-NOV-1999; 99US-0161992.
PR 02-NOV-1999; 99US-0161993.
PR 03-NOV-1999; 99US-0162142.
PR 04-NOV-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1-266-42
Score: 222.00
Percent Similarity: 100.00%
Best Local Similarity: 95.50%
Query Match: 100.00%
DB: 21

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC49242 (1-771)

QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCGGATGGATCGTACGATCTACGTTGGGATTTGCTGGAGATATTCGCGAG 60

QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
DB 61 TGTGAGGTTGAAGATCTCTTACAGATATGACCAATTTGGACATTTGACATTTGAAGATT 120

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCCCTGGTTATGCCCTTTCGAGTTTGAAGATCCCTCGTGATCGACAGAT 180

QY 61 AlaIleTyrGlyValGlyAspGlyTyrAspPheAspGlyCysAtgLeuArgValGluIleAla 80
DB 181 GCAATTATGACGATGATGATGATTTGATGGGTGTCGACTTCGGGTTGAGATTGCA 240

QY 81 HisGlyGlyValArgArg*****SerTyrSerAlaSer 100
DB 241 CATGTGTGTCGTAGATTTTCCATCAGTTGATAGTACGACGACGACGCTACAGTGCAGC 300

QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACACCTTCAAGACGCTCTGACTACCGCTGCTTGTGACCGGATTTACCGCCTTCTCT 360
```

```
QY 121 SerTIPGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGGCAGGACCTTAAGGATCATCGCGCAAGCTGGAGATGCTGCTTCTCTGAAGTT 420

QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB 421 TTCCTGACCGTAAAGGCATGTCGGGGTTGGATTATAGCAACTATGATGATATGAAG 480

QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TAGCAATAAGGAACTTGATGCCACTGAATTTGGAATCTTCTCTAGTCTTATATA 540

QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGTGAGGAATATGATGTCGAGAGTGTGAGTCGAGCCAGATGATTTCTAAAGCTAT 600

QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
DB 601 AGAAGCAGGAGTCGGAGCCGCTGGTCCCAAGCTGTAGCTATATAGTAGCAAGACGAGGTG 660

QY 221 SerPro 222
DB 661 TCACCT 666

RESULT 2
AAC49248
ID AAC49248 standard; DNA; 925 BP.
AC AAC49248;
XX
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 60471.
DE Arabidopsis thaliana
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
```

PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135363.  
PR 24-MAY-1999; 99US-0135623.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137503.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 11-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139115.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140693.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142053.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0161404.  
PR 22-OCT-1999; 99US-0161405.

```

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,398-42 Length: 925
Score: 222.00 Matches: 212
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.58% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC49248 (1-925)

QY 1 MetSerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgIys 20
Db 134 ATGAGTAGCGGATGAATCGTACGATCTACGTTGGGAATTCCTCGGAGATATTCGCAAG 193
QY 21 CysGluValGluAspLeuPheTyrIysTyrGlyProIleValAspIleAspLeuIysIle 40
Db 194 TGTGAGGTTGAGATCTCTTCTACAGTATGACCAATGTGGACATTTGAAGATT 253
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 254 CCACCGAGACCTCCCTGGTTATGCCCTTTCGAGTTTGAGATCCTCGTGATCGAGCAT 313
QY 61 AlalleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 314 GCAATTTATGACGCTGATGTTATGATTTGATGGTGTGCGATTCGGGTTGAGATTGCA 373
QY 81 HisGlyGlyArgArg******SeryTyrSerAlaser 100
Db 374 CATGTGGTCGAGATTTTCCACATCAGTTGATAGGTACGACGACGACGATACAGTGGAGC 433
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
Db 434 CGTGACCTTCAAGACGCTCTGACTACCGCGTGTGACCGGATTTACCGGATTTCTGCT 493
QY 121 SerTyrGlnAspLeuAspHisMetArgIysAlaGlyAspValCysPheSerGluVal 140
Db 494 TGTGTGCAGGACCTTAAGCATCACATGCGCAAGCTGGAGATGCTGCTTCTCTGAAGTT 553
QY 141 PheProAspArgGlyGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetIys 160
Db 554 TTCCCTGACCGTAAAGGATGCTGGGTTGTGGATTATAGCAACTATGATGATGAAG 613
QY 161 TyrAlaIleArgIysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 614 TACGCAATAAGAAACTTGAATCCACTGAATTCGAAATGCTTTCTCTAGTGGCTTATATA 673
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerIysSerTyr 200
Db 674 CGGGTGAGGGAATATAGTCGAGGAGTGATGTCGAAGCCAGATGATCTTAAGGCTAT 733
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerIysSerArgSerVal 220
Db 734 AGAAGCAGGATCGGAGCGCTGCTCCAGCTGTAGCTATAGTAGCAAGCAGGAGGTGTG 793
QY 221 SerPro 222
Db 794 TCACCT 799

RESULT 3
AAC40180
ID AAC40180 standard; DNA; 1213 BP.
XX
AC AAC40180;
XX
```

```

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 27352.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; terminator sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
```



|    |              |               |                                                            |              |                                                                      |
|----|--------------|---------------|------------------------------------------------------------|--------------|----------------------------------------------------------------------|
| PR | 18-JUN-1999; | 99US-0139763. | PR                                                         | 27-AUG-1999; | 99US-0151066.                                                        |
| PR | 21-JUN-1999; | 99US-0139817. | PR                                                         | 27-AUG-1999; | 99US-0151080.                                                        |
| PR | 22-JUN-1999; | 99US-0139899. | PR                                                         | 30-AUG-1999; | 99US-0151303.                                                        |
| PR | 23-JUN-1999; | 99US-0140353. | PR                                                         | 31-AUG-1999; | 99US-0151438.                                                        |
| PR | 23-JUN-1999; | 99US-0140354. | PR                                                         | 01-SEP-1999; | 99US-0151930.                                                        |
| PR | 24-JUN-1999; | 99US-0140695. | PR                                                         | 07-SEP-1999; | 99US-0152363.                                                        |
| PR | 28-JUN-1999; | 99US-0140823. | PR                                                         | 10-SEP-1999; | 99US-0153070.                                                        |
| PR | 29-JUN-1999; | 99US-0140931. | PR                                                         | 13-SEP-1999; | 99US-0153758.                                                        |
| PR | 30-JUN-1999; | 99US-0141287. | PR                                                         | 16-SEP-1999; | 99US-0154018.                                                        |
| PR | 01-JUL-1999; | 99US-0141842. | PR                                                         | 16-SEP-1999; | 99US-0154039.                                                        |
| PR | 01-JUL-1999; | 99US-0142154. | PR                                                         | 20-SEP-1999; | 99US-0154779.                                                        |
| PR | 02-JUL-1999; | 99US-0142055. | PR                                                         | 22-SEP-1999; | 99US-0155139.                                                        |
| PR | 06-JUL-1999; | 99US-0142390. | PR                                                         | 23-SEP-1999; | 99US-0155486.                                                        |
| PR | 08-JUL-1999; | 99US-0142803. | PR                                                         | 24-SEP-1999; | 99US-0155659.                                                        |
| PR | 09-JUL-1999; | 99US-0143920. | PR                                                         | 28-SEP-1999; | 99US-0156458.                                                        |
| PR | 12-JUL-1999; | 99US-0143977. | PR                                                         | 29-SEP-1999; | 99US-0158596.                                                        |
| PR | 13-JUL-1999; | 99US-0143542. | PR                                                         | 04-OCT-1999; | 99US-0157117.                                                        |
| PR | 14-JUL-1999; | 99US-0143624. | PR                                                         | 05-OCT-1999; | 99US-0157753.                                                        |
| PR | 15-JUL-1999; | 99US-0144005. | PR                                                         | 06-OCT-1999; | 99US-0157865.                                                        |
| PR | 16-JUL-1999; | 99US-0144085. | PR                                                         | 07-OCT-1999; | 99US-0158029.                                                        |
| PR | 16-JUL-1999; | 99US-0144086. | PR                                                         | 08-OCT-1999; | 99US-0158232.                                                        |
| PR | 19-JUL-1999; | 99US-0144325. | PR                                                         | 12-OCT-1999; | 99US-0158369.                                                        |
| PR | 19-JUL-1999; | 99US-0144331. | PR                                                         | 13-OCT-1999; | 99US-0159293.                                                        |
| PR | 19-JUL-1999; | 99US-0144332. | PR                                                         | 13-OCT-1999; | 99US-0159294.                                                        |
| PR | 19-JUL-1999; | 99US-0144333. | PR                                                         | 13-OCT-1999; | 99US-0159295.                                                        |
| PR | 19-JUL-1999; | 99US-0144334. | PR                                                         | 14-OCT-1999; | 99US-0159329.                                                        |
| PR | 19-JUL-1999; | 99US-0144335. | PR                                                         | 14-OCT-1999; | 99US-0159330.                                                        |
| PR | 20-JUL-1999; | 99US-0144352. | PR                                                         | 14-OCT-1999; | 99US-0159331.                                                        |
| PR | 20-JUL-1999; | 99US-0144632. | PR                                                         | 14-OCT-1999; | 99US-0159637.                                                        |
| PR | 20-JUL-1999; | 99US-0144884. | PR                                                         | 14-OCT-1999; | 99US-0159638.                                                        |
| PR | 21-JUL-1999; | 99US-0144814. | PR                                                         | 18-OCT-1999; | 99US-0159584.                                                        |
| PR | 21-JUL-1999; | 99US-0145086. | PR                                                         | 21-OCT-1999; | 99US-0160741.                                                        |
| PR | 21-JUL-1999; | 99US-0145088. | PR                                                         | 21-OCT-1999; | 99US-0160767.                                                        |
| PR | 22-JUL-1999; | 99US-0145085. | PR                                                         | 21-OCT-1999; | 99US-0160768.                                                        |
| PR | 22-JUL-1999; | 99US-0145087. | PR                                                         | 21-OCT-1999; | 99US-0160770.                                                        |
| PR | 22-JUL-1999; | 99US-0145089. | PR                                                         | 21-OCT-1999; | 99US-0160814.                                                        |
| PR | 22-JUL-1999; | 99US-0145192. | PR                                                         | 21-OCT-1999; | 99US-0160815.                                                        |
| PR | 23-JUL-1999; | 99US-0145145. | PR                                                         | 22-OCT-1999; | 99US-0160980.                                                        |
| PR | 23-JUL-1999; | 99US-0145218. | PR                                                         | 22-OCT-1999; | 99US-0160981.                                                        |
| PR | 23-JUL-1999; | 99US-0145224. | PR                                                         | 22-OCT-1999; | 99US-0160989.                                                        |
| PR | 26-JUL-1999; | 99US-0145276. | PR                                                         | 25-OCT-1999; | 99US-0161404.                                                        |
| PR | 27-JUL-1999; | 99US-0145913. | PR                                                         | 25-OCT-1999; | 99US-0161405.                                                        |
| PR | 27-JUL-1999; | 99US-0145918. | PR                                                         | 25-OCT-1999; | 99US-0161406.                                                        |
| PR | 27-JUL-1999; | 99US-0145919. | PR                                                         | 26-OCT-1999; | 99US-0161359.                                                        |
| PR | 28-JUL-1999; | 99US-0145951. | PR                                                         | 26-OCT-1999; | 99US-0161360.                                                        |
| PR | 02-AUG-1999; | 99US-0146385. | PR                                                         | 26-OCT-1999; | 99US-0161361.                                                        |
| PR | 02-AUG-1999; | 99US-0146388. | PR                                                         | 28-OCT-1999; | 99US-0161920.                                                        |
| PR | 02-AUG-1999; | 99US-0146389. | PR                                                         | 28-OCT-1999; | 99US-0161992.                                                        |
| PR | 03-AUG-1999; | 99US-0147038. | PR                                                         | 28-OCT-1999; | 99US-0161993.                                                        |
| PR | 04-AUG-1999; | 99US-0147204. | PR                                                         | 29-OCT-1999; | 99US-0162142.                                                        |
| PR | 05-AUG-1999; | 99US-0147302. |                                                            |              |                                                                      |
| PR | 05-AUG-1999; | 99US-0147192. | Alignment Scores:                                          |              |                                                                      |
| PR | 05-AUG-1999; | 99US-0147260. | Pred. No.:                                                 | 6.18e-42     | 1213                                                                 |
| PR | 06-AUG-1999; | 99US-0147303. | Score:                                                     | 222.00       | 212                                                                  |
| PR | 06-AUG-1999; | 99US-0147416. | Percent Similarity:                                        | 100.00%      | Matches: 10                                                          |
| PR | 09-AUG-1999; | 99US-0147493. | Best Local Similarity:                                     | 95.50%       | Mismatches: 0                                                        |
| PR | 09-AUG-1999; | 99US-0147935. | Query Match:                                               | 100.00%      | Indels: 0                                                            |
| PR | 10-AUG-1999; | 99US-0148171. | DB:                                                        | 21           | Gaps: 0                                                              |
| PR | 11-AUG-1999; | 99US-0148319. |                                                            |              |                                                                      |
| PR | 12-AUG-1999; | 99US-0148341. | US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC40180 (1-1213) |              |                                                                      |
| PR | 13-AUG-1999; | 99US-0148565. | QY                                                         | 1            | MetSerSerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20      |
| PR | 13-AUG-1999; | 99US-0148684. | Db                                                         | 134          | ATGAGTAGCCGATGGAATCGTAGCATCTCTTACAGTATGGACCAATTGGACATTGTTGAAGATT 193 |
| PR | 16-AUG-1999; | 99US-0149368. |                                                            |              |                                                                      |
| PR | 17-AUG-1999; | 99US-0149375. | QY                                                         | 21           | CysGluValGluAspLeuPheTyrIleValAspIleAspLeuLysIle 40                  |
| PR | 18-AUG-1999; | 99US-0149426. | Db                                                         | 194          | TGTGAGTTGAAGATCTCTTCTACAGTATGGACCAATTGGACATTGTTGAAGATT 253           |
| PR | 20-AUG-1999; | 99US-0149722. |                                                            |              |                                                                      |
| PR | 20-AUG-1999; | 99US-0149723. | QY                                                         | 41           | ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60            |
| PR | 20-AUG-1999; | 99US-0149929. | Db                                                         | 254          | CCACCGAGACCTCTCGTTATGCCCTTTCGAGTTTGAAGATCCTCTCGTATGCACACGAT 313      |
| PR | 23-AUG-1999; | 99US-0149902. |                                                            |              |                                                                      |
| PR | 23-AUG-1999; | 99US-0149930. |                                                            |              |                                                                      |
| PR | 25-AUG-1999; | 99US-0150565. |                                                            |              |                                                                      |
| PR | 26-AUG-1999; | 99US-0150884. |                                                            |              |                                                                      |
| PR | 27-AUG-1999; | 99US-0151065. |                                                            |              |                                                                      |

|    |              |               |
|----|--------------|---------------|
| PR | 30-APR-1999; | 99US-01320448 |
| PR | 30-APR-1999; | 99US-01324407 |
| PR | 04-MAY-1999; | 99US-01324484 |
| PR | 05-MAY-1999; | 99US-01324884 |
| PR | 06-MAY-1999; | 99US-01324886 |
| PR | 06-MAY-1999; | 99US-01324887 |
| PR | 07-MAY-1999; | 99US-01328663 |
| PR | 11-MAY-1999; | 99US-01324256 |
| PR | 14-MAY-1999; | 99US-01324218 |
| PR | 14-MAY-1999; | 99US-01324219 |
| PR | 14-MAY-1999; | 99US-01324221 |
| PR | 14-MAY-1999; | 99US-01343370 |
| PR | 18-MAY-1999; | 99US-01347668 |
| PR | 19-MAY-1999; | 99US-01349411 |
| PR | 20-MAY-1999; | 99US-01353524 |
| PR | 21-MAY-1999; | 99US-01353543 |
| PR | 24-MAY-1999; | 99US-01356629 |
| PR | 25-MAY-1999; | 99US-01360211 |
| PR | 27-MAY-1999; | 99US-01363992 |
| PR | 28-MAY-1999; | 99US-01367862 |
| PR | 01-JUN-1999; | 99US-01372822 |
| PR | 03-JUN-1999; | 99US-01375288 |
| PR | 04-JUN-1999; | 99US-01375002 |
| PR | 07-JUN-1999; | 99US-01377244 |
| PR | 08-JUN-1999; | 99US-01380694 |
| PR | 10-JUN-1999; | 99US-01385400 |
| PR | 10-JUN-1999; | 99US-01388847 |
| PR | 14-JUN-1999; | 99US-01391149 |
| PR | 16-JUN-1999; | 99US-01394522 |
| PR | 16-JUN-1999; | 99US-01394532 |
| PR | 17-JUN-1999; | 99US-01394592 |
| PR | 18-JUN-1999; | 99US-01394555 |
| PR | 18-JUN-1999; | 99US-01394555 |
| PR | 18-JUN-1999; | 99US-01394566 |
| PR | 18-JUN-1999; | 99US-01394578 |
| PR | 18-JUN-1999; | 99US-01394588 |
| PR | 18-JUN-1999; | 99US-01394600 |
| PR | 18-JUN-1999; | 99US-01394611 |
| PR | 18-JUN-1999; | 99US-01394623 |
| PR | 18-JUN-1999; | 99US-01394636 |
| PR | 18-JUN-1999; | 99US-01397500 |
| PR | 18-JUN-1999; | 99US-01397563 |
| PR | 21-JUN-1999; | 99US-01398177 |
| PR | 22-JUN-1999; | 99US-01398699 |
| PR | 23-JUN-1999; | 99US-01403553 |
| PR | 23-JUN-1999; | 99US-01403554 |
| PR | 24-JUN-1999; | 99US-01406695 |
| PR | 24-JUN-1999; | 99US-01406955 |
| PR | 28-JUN-1999; | 99US-01408223 |
| PR | 29-JUN-1999; | 99US-01409991 |
| PR | 30-JUN-1999; | 99US-01412887 |
| PR | 01-JUL-1999; | 99US-01418442 |
| PR | 01-JUL-1999; | 99US-01421544 |
| PR | 02-JUL-1999; | 99US-01420555 |
| PR | 06-JUL-1999; | 99US-01423900 |
| PR | 08-JUL-1999; | 99US-01428003 |
| PR | 09-JUL-1999; | 99US-01429200 |
| PR | 12-JUL-1999; | 99US-01429977 |
| PR | 13-JUL-1999; | 99US-01435442 |
| PR | 14-JUL-1999; | 99US-01436622 |
| PR | 15-JUL-1999; | 99US-01440005 |
| PR | 16-JUL-1999; | 99US-01440085 |
| PR | 16-JUL-1999; | 99US-01440886 |
| PR | 19-JUL-1999; | 99US-01443225 |
| PR | 19-JUL-1999; | 99US-01443311 |
| PR | 19-JUL-1999; | 99US-01443332 |
| PR | 19-JUL-1999; | 99US-01443333 |
| PR | 19-JUL-1999; | 99US-01443334 |
| PR | 19-JUL-1999; | 99US-01443335 |
| PR | 20-JUL-1999; | 99US-01443352 |
| PR | 20-JUL-1999; | 99US-01446632 |
| PR | 20-JUL-1999; | 99US-01448814 |
| PR | 21-JUL-1999; | 99US-01448144 |

|                        |                                                                  |               |                                                           |                                                                  |               |
|------------------------|------------------------------------------------------------------|---------------|-----------------------------------------------------------|------------------------------------------------------------------|---------------|
| PR                     | 21-JUL-1999;                                                     | 99US-0145086. | PR                                                        | 21-OCT-1999;                                                     | 99US-0160741. |
| PR                     | 21-JUL-1999;                                                     | 99US-0145088. | PR                                                        | 21-OCT-1999;                                                     | 99US-0160767. |
| PR                     | 22-JUL-1999;                                                     | 99US-0145085. | PR                                                        | 21-OCT-1999;                                                     | 99US-0160768. |
| PR                     | 22-JUL-1999;                                                     | 99US-0145087. | PR                                                        | 21-OCT-1999;                                                     | 99US-0160770. |
| PR                     | 22-JUL-1999;                                                     | 99US-0145089. | PR                                                        | 21-OCT-1999;                                                     | 99US-0160814. |
| PR                     | 22-JUL-1999;                                                     | 99US-0145192. | PR                                                        | 21-OCT-1999;                                                     | 99US-0160815. |
| PR                     | 23-JUL-1999;                                                     | 99US-0145119. | PR                                                        | 22-OCT-1999;                                                     | 99US-0160980. |
| PR                     | 23-JUL-1999;                                                     | 99US-0145218. | PR                                                        | 22-OCT-1999;                                                     | 99US-0160981. |
| PR                     | 23-JUL-1999;                                                     | 99US-0145224. | PR                                                        | 22-OCT-1999;                                                     | 99US-0160989. |
| PR                     | 26-JUL-1999;                                                     | 99US-0145276. | PR                                                        | 23-OCT-1999;                                                     | 99US-0161404. |
| PR                     | 27-JUL-1999;                                                     | 99US-0145913. | PR                                                        | 23-OCT-1999;                                                     | 99US-0161405. |
| PR                     | 27-JUL-1999;                                                     | 99US-0145918. | PR                                                        | 25-OCT-1999;                                                     | 99US-0161406. |
| PR                     | 27-JUL-1999;                                                     | 99US-0145919. | PR                                                        | 26-OCT-1999;                                                     | 99US-0161359. |
| PR                     | 28-JUL-1999;                                                     | 99US-0145951. | PR                                                        | 26-OCT-1999;                                                     | 99US-0161360. |
| PR                     | 02-AUG-1999;                                                     | 99US-0146386. | PR                                                        | 26-OCT-1999;                                                     | 99US-0161361. |
| PR                     | 02-AUG-1999;                                                     | 99US-0146388. | PR                                                        | 28-OCT-1999;                                                     | 99US-0161920. |
| PR                     | 03-AUG-1999;                                                     | 99US-0146389. | PR                                                        | 28-OCT-1999;                                                     | 99US-0161992. |
| PR                     | 03-AUG-1999;                                                     | 99US-0147038. | PR                                                        | 28-OCT-1999;                                                     | 99US-0161993. |
| PR                     | 04-AUG-1999;                                                     | 99US-0147204. | PR                                                        | 29-OCT-1999;                                                     | 99US-0162142. |
| PR                     | 04-AUG-1999;                                                     | 99US-0147302. |                                                           |                                                                  |               |
| PR                     | 05-AUG-1999;                                                     | 99US-0147192. |                                                           |                                                                  |               |
| PR                     | 05-AUG-1999;                                                     | 99US-0147260. |                                                           |                                                                  |               |
| PR                     | 06-AUG-1999;                                                     | 99US-0147303. |                                                           |                                                                  |               |
| PR                     | 06-AUG-1999;                                                     | 99US-0147416. |                                                           |                                                                  |               |
| PR                     | 09-AUG-1999;                                                     | 99US-0147493. |                                                           |                                                                  |               |
| PR                     | 09-AUG-1999;                                                     | 99US-0147935. |                                                           |                                                                  |               |
| PR                     | 10-AUG-1999;                                                     | 99US-0148171. |                                                           |                                                                  |               |
| PR                     | 11-AUG-1999;                                                     | 99US-0148319. |                                                           |                                                                  |               |
| PR                     | 12-AUG-1999;                                                     | 99US-0148341. |                                                           |                                                                  |               |
| PR                     | 13-AUG-1999;                                                     | 99US-0148565. |                                                           |                                                                  |               |
| PR                     | 13-AUG-1999;                                                     | 99US-0148684. |                                                           |                                                                  |               |
| PR                     | 16-AUG-1999;                                                     | 99US-0149368. |                                                           |                                                                  |               |
| PR                     | 17-AUG-1999;                                                     | 99US-0149175. |                                                           |                                                                  |               |
| PR                     | 18-AUG-1999;                                                     | 99US-0149426. |                                                           |                                                                  |               |
| PR                     | 20-AUG-1999;                                                     | 99US-0149722. |                                                           |                                                                  |               |
| PR                     | 20-AUG-1999;                                                     | 99US-0149723. |                                                           |                                                                  |               |
| PR                     | 20-AUG-1999;                                                     | 99US-0149929. |                                                           |                                                                  |               |
| PR                     | 23-AUG-1999;                                                     | 99US-0149902. |                                                           |                                                                  |               |
| PR                     | 23-AUG-1999;                                                     | 99US-0149930. |                                                           |                                                                  |               |
| PR                     | 25-AUG-1999;                                                     | 99US-0150566. |                                                           |                                                                  |               |
| PR                     | 26-AUG-1999;                                                     | 99US-0150884. |                                                           |                                                                  |               |
| PR                     | 27-AUG-1999;                                                     | 99US-0151065. |                                                           |                                                                  |               |
| PR                     | 27-AUG-1999;                                                     | 99US-0151066. |                                                           |                                                                  |               |
| PR                     | 27-AUG-1999;                                                     | 99US-0151080. |                                                           |                                                                  |               |
| PR                     | 30-AUG-1999;                                                     | 99US-0151303. |                                                           |                                                                  |               |
| PR                     | 31-AUG-1999;                                                     | 99US-0151438. |                                                           |                                                                  |               |
| PR                     | 01-SEP-1999;                                                     | 99US-0151930. |                                                           |                                                                  |               |
| PR                     | 07-SEP-1999;                                                     | 99US-0152363. |                                                           |                                                                  |               |
| PR                     | 10-SEP-1999;                                                     | 99US-0153070. |                                                           |                                                                  |               |
| PR                     | 13-SEP-1999;                                                     | 99US-0153758. |                                                           |                                                                  |               |
| PR                     | 15-SEP-1999;                                                     | 99US-0154018. |                                                           |                                                                  |               |
| PR                     | 16-SEP-1999;                                                     | 99US-0154039. |                                                           |                                                                  |               |
| PR                     | 20-SEP-1999;                                                     | 99US-0154779. |                                                           |                                                                  |               |
| PR                     | 22-SEP-1999;                                                     | 99US-0155139. |                                                           |                                                                  |               |
| PR                     | 23-SEP-1999;                                                     | 99US-0155486. |                                                           |                                                                  |               |
| PR                     | 24-SEP-1999;                                                     | 99US-0155659. |                                                           |                                                                  |               |
| PR                     | 28-SEP-1999;                                                     | 99US-0156458. |                                                           |                                                                  |               |
| PR                     | 29-SEP-1999;                                                     | 99US-0156596. |                                                           |                                                                  |               |
| PR                     | 04-OCT-1999;                                                     | 99US-0157117. |                                                           |                                                                  |               |
| PR                     | 05-OCT-1999;                                                     | 99US-0157753. |                                                           |                                                                  |               |
| PR                     | 06-OCT-1999;                                                     | 99US-0157865. |                                                           |                                                                  |               |
| PR                     | 08-OCT-1999;                                                     | 99US-0158029. |                                                           |                                                                  |               |
| PR                     | 08-OCT-1999;                                                     | 99US-0158232. |                                                           |                                                                  |               |
| PR                     | 12-OCT-1999;                                                     | 99US-0158369. |                                                           |                                                                  |               |
| PR                     | 13-OCT-1999;                                                     | 99US-0159293. |                                                           |                                                                  |               |
| PR                     | 13-OCT-1999;                                                     | 99US-0159294. |                                                           |                                                                  |               |
| PR                     | 13-OCT-1999;                                                     | 99US-0159295. |                                                           |                                                                  |               |
| PR                     | 14-OCT-1999;                                                     | 99US-0159329. |                                                           |                                                                  |               |
| PR                     | 14-OCT-1999;                                                     | 99US-0159330. |                                                           |                                                                  |               |
| PR                     | 14-OCT-1999;                                                     | 99US-0159331. |                                                           |                                                                  |               |
| PR                     | 14-OCT-1999;                                                     | 99US-0159637. |                                                           |                                                                  |               |
| PR                     | 14-OCT-1999;                                                     | 99US-0159638. |                                                           |                                                                  |               |
| PR                     | 18-OCT-1999;                                                     | 99US-0159584. |                                                           |                                                                  |               |
| Alignment Scores:      |                                                                  |               | US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC42865 (1-714) |                                                                  |               |
| Pred. No.:             |                                                                  |               | 9.64e-42                                                  |                                                                  |               |
| Score:                 |                                                                  |               | 218.00                                                    |                                                                  |               |
| Percent Similarity:    |                                                                  |               | 100.00%                                                   |                                                                  |               |
| Best Local Similarity: |                                                                  |               | 95.41%                                                    |                                                                  |               |
| Query Match:           |                                                                  |               | 96.20%                                                    |                                                                  |               |
| DB:                    |                                                                  |               | 21                                                        |                                                                  |               |
| Length:                |                                                                  |               | 714                                                       |                                                                  |               |
| Matches:               |                                                                  |               | 208                                                       |                                                                  |               |
| Conservative:          |                                                                  |               | 10                                                        |                                                                  |               |
| Mismatch:              |                                                                  |               | 0                                                         |                                                                  |               |
| Indels:                |                                                                  |               | 0                                                         |                                                                  |               |
| Gaps:                  |                                                                  |               | 0                                                         |                                                                  |               |
| QY                     | 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   | 20            | QY                                                        | 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   | 20            |
| DB                     | 1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGAAATTCCTCGGAGATATTCGCAAG    | 60            | DB                                                        | 1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGAAATTCCTCGGAGATATTCGCAAG    | 60            |
| QY                     | 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle  | 40            | QY                                                        | 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle  | 40            |
| DB                     | 61 TGTGAGTTGAGATCTCTTCTACAGATATGACCAATTTGGACATTTGAGATTGAGATT     | 120           | DB                                                        | 61 TGTGAGTTGAGATCTCTTCTACAGATATGACCAATTTGGACATTTGAGATTGAGATT     | 120           |
| QY                     | 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp  | 60            | QY                                                        | 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp  | 60            |
| DB                     | 121 CCACCGAGACCTCTCGTTATGTCCTTCGAGTTTGAAGATCCTCGTGATCGACAGCAT    | 180           | DB                                                        | 121 CCACCGAGACCTCTCGTTATGTCCTTCGAGTTTGAAGATCCTCGTGATCGACAGCAT    | 180           |
| QY                     | 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla  | 80            | QY                                                        | 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla  | 80            |
| DB                     | 181 GCAATTTATGACGCTGATGTTGATTTGATGGGTGTCGACTTCGGGTTGAGATTGCA     | 240           | DB                                                        | 181 GCAATTTATGACGCTGATGTTGATTTGATGGGTGTCGACTTCGGGTTGAGATTGCA     | 240           |
| QY                     | 81 HisGlyGlyArgArgG                                              | 100           | QY                                                        | 81 HisGlyGlyArgArgG                                              | 100           |
| DB                     | 241 CATGTTGTTGTTAGATTTTCACCATCAGTTAGGTACAGCAGCAGCTAGTCCGAGC      | 300           | DB                                                        | 241 CATGTTGTTGTTAGATTTTCACCATCAGTTAGGTACAGCAGCAGCTAGTCCGAGC      | 300           |
| QY                     | 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla | 120           | QY                                                        | 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla | 120           |
| DB                     | 301 CGTGACCTTCAGACGCTCTGACTACCGCGTCTTGTGACCGGATTAACCGCTCTCT      | 360           | DB                                                        | 301 CGTGACCTTCAGACGCTCTGACTACCGCGTCTTGTGACCGGATTAACCGCTCTCT      | 360           |
| QY                     | 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal | 140           | QY                                                        | 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal | 140           |
| DB                     | 361 TCGTGGCAGGACCTTAGGATCACATGGCGAAGCTGGAGATGCTGCTTCTCTGAATT     | 420           | DB                                                        | 361 TCGTGGCAGGACCTTAGGATCACATGGCGAAGCTGGAGATGCTGCTTCTCTGAATT     | 420           |
| QY                     | 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys | 160           | QY                                                        | 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys | 160           |
| DB                     | 421 TTCCCTGACCGTAAAGGATCTCTGGGTTGTGATTTAGCAACTATGATGATGATGAAG    | 480           | DB                                                        | 421 TTCCCTGACCGTAAAGGATCTCTGGGTTGTGATTTAGCAACTATGATGATGATGAAG    | 480           |
| QY                     | 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle | 180           | QY                                                        | 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle | 180           |
| DB                     | 481 TACGCAATAGAAACTTGTATGTCACCTGAAATTTGAAATGCTTCTCTAGTCTTATA     | 540           | DB                                                        | 481 TACGCAATAGAAACTTGTATGTCACCTGAAATTTGAAATGCTTCTCTAGTCTTATA     | 540           |
| QY                     | 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr | 200           | QY                                                        | 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr | 200           |
| DB                     | 541 CGGGTGAGGGAATATGATGCTGAGGAGTGTGAGTCGAGCCCGAGATGATTTCTAAAGCAT | 600           | DB                                                        | 541 CGGGTGAGGGAATATGATGCTGAGGAGTGTGAGTCGAGCCCGAGATGATTTCTAAAGCAT | 600           |
| QY                     | 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArg       | 218           | QY                                                        | 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArg       | 218           |
| DB                     | 601 AGAGCAGGAGTCGAGCGCTGTCCTCAAGCTGTAGCTATAGTAGCAAGAGCAGG        | 654           | DB                                                        | 601 AGAGCAGGAGTCGAGCGCTGTCCTCAAGCTGTAGCTATAGTAGCAAGAGCAGG        | 654           |



PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132486.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-014218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139432.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147933.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.

|                                                            |              |                                                              |      |
|------------------------------------------------------------|--------------|--------------------------------------------------------------|------|
| PR                                                         | 21-OCT-1999; | 99US-0160741.                                                |      |
| PR                                                         | 21-OCT-1999; | 99US-0160767.                                                |      |
| PR                                                         | 21-OCT-1999; | 99US-0160768.                                                |      |
| PR                                                         | 21-OCT-1999; | 99US-0160770.                                                |      |
| PR                                                         | 21-OCT-1999; | 99US-0160814.                                                |      |
| PR                                                         | 21-OCT-1999; | 99US-0160815.                                                |      |
| PR                                                         | 22-OCT-1999; | 99US-0160980.                                                |      |
| PR                                                         | 22-OCT-1999; | 99US-0160981.                                                |      |
| PR                                                         | 22-OCT-1999; | 99US-0160989.                                                |      |
| PR                                                         | 23-OCT-1999; | 99US-0161404.                                                |      |
| PR                                                         | 23-OCT-1999; | 99US-0161405.                                                |      |
| PR                                                         | 25-OCT-1999; | 99US-0161406.                                                |      |
| PR                                                         | 26-OCT-1999; | 99US-0161359.                                                |      |
| PR                                                         | 26-OCT-1999; | 99US-0161360.                                                |      |
| PR                                                         | 28-OCT-1999; | 99US-0161361.                                                |      |
| PR                                                         | 28-OCT-1999; | 99US-0161920.                                                |      |
| PR                                                         | 28-OCT-1999; | 99US-0161992.                                                |      |
| PR                                                         | 28-OCT-1999; | 99US-0161993.                                                |      |
| PR                                                         | 29-OCT-1999; | 99US-0162142.                                                |      |
| Alignment Scores:                                          |              |                                                              |      |
| Pred. No.:                                                 | 3, 87e-22    | Length:                                                      | 1232 |
| Score:                                                     | 143.00       | Matches:                                                     | 148  |
| Percent Similarity:                                        | 68.40%       | Conservative:                                                | 10   |
| Best Local Similarity:                                     | 64.07%       | Mismatches:                                                  | 63   |
| Query Match:                                               | 64.41%       | Indels:                                                      | 10   |
| DB:                                                        | 21           | Gaps:                                                        | 1    |
| US-10-014-927-19MOD_COPY_1_222 (1-222) x AAC34814 (1-1232) |              |                                                              |      |
| QY                                                         | 1            | MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys | 20   |
| DB                                                         | 94           | ATGAGCAGTCTGTCGAGTAGAACCGTGTACGTCGGAACCTCTCGCGATATCCGTGAG    | 153  |
| QY                                                         | 21           | CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle | 40   |
| DB                                                         | 154          | AGAGAGTCCAGATTGTTTCAGTAGTAGTACCTGTTGTTCAAATTGATTGAGGTT       | 213  |
| QY                                                         | 41           | ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp | 60   |
| DB                                                         | 214          | CTCCAGAGCTCTCGTGTATGATTCGTTGATTTGATGATGCTCGGATGCTGAGAT       | 273  |
| QY                                                         | 61           | AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla | 80   |
| DB                                                         | 274          | GCTATTATGCTGATGCTATGACTTTGATGGGCATCGTTGAGGCTGGAATTGGCG       | 333  |
| QY                                                         | 81           | HisGlyGlyArgArg-----                                         | 90   |
| DB                                                         | 334          | CATGGTGGAGGCGTTTCATCAGATGATCTCGGGTAGTTTCAATGCTGGTGGCGGTGT    | 393  |
| QY                                                         | 91           | *****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg           | 110  |
| DB                                                         | 394          | GGTGGTGGCGGCGTGGTGTGATGTTGTTGGGCCATCTAGGAGATCAGAGTTTCGT      | 453  |
| QY                                                         | 111          | ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg | 130  |
| DB                                                         | 454          | GTTCATGTCACAGCTTCCTTCATCTGCTTCTTGGCAAGATCTCAAGATCACATGCT     | 513  |
| QY                                                         | 131          | LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal | 150  |
| DB                                                         | 514          | AAAGGAGCGCATGCTGTTCTCGCAAGTGTACCGTGATGCTAGAGGACCACTGGAGTT    | 573  |
| QY                                                         | 151          | ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu | 170  |
| DB                                                         | 574          | GTGTATTACACCTGCTATGAGACATGATGTCGCTGAAAAGCTCAGACACAGAG        | 633  |
| QY                                                         | 171          | PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal | 190  |
| DB                                                         | 634          | TTTCGAATCGTTTTCGATGATGATATGTCGGGTTAGAGATATGATTCAGGAGAGAT     | 693  |
| QY                                                         | 191          | SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGlyProSer       | 210  |
| DB                                                         | 694          | TCTAGGAGTCTAGCGGGGAAGATCTATTCTTAAGAGCGGACGCGGCGGTGGACGA      | 753  |

|          |                                                                |                                   |     |
|----------|----------------------------------------------------------------|-----------------------------------|-----|
| QY       | 211                                                            | CysSerTyrSerSerLysSerArgSerValSer | 221 |
| DB       | 754                                                            | ACCGTGAGCCGAGCAGGAGCAGAGCAGGAGC   | 786 |
| RESULT 7 |                                                                |                                   |     |
| AAC38528 |                                                                |                                   |     |
| ID       | AAC38528                                                       | standard; DNA; 1190 BP.           |     |
| XX       |                                                                |                                   |     |
| AC       | AAC38528;                                                      |                                   |     |
| XX       |                                                                |                                   |     |
| DT       | 17-OCT-2000                                                    | (first entry)                     |     |
| XX       |                                                                |                                   |     |
| DE       | Arabidopsis thaliana                                           | DNA fragment SEQ ID NO: 21293.    |     |
| XX       |                                                                |                                   |     |
| KW       | Hybridisation assay; Genetic mapping; gene expression control; |                                   |     |
| KW       | protein identification; signal transduction pathway;           |                                   |     |
| KW       | metabolic pathway; promoter; termination sequence; ss.         |                                   |     |
| XX       |                                                                |                                   |     |
| OS       | Arabidopsis thaliana.                                          |                                   |     |
| XX       |                                                                |                                   |     |
| PN       | EP1033405-A2.                                                  |                                   |     |
| XX       |                                                                |                                   |     |
| PD       | 06-SEP-2000.                                                   |                                   |     |
| XX       |                                                                |                                   |     |
| PF       | 25-FEB-2000; 2000EP-0301439.                                   |                                   |     |
| XX       |                                                                |                                   |     |
| PR       | 25-FEB-1999;                                                   | 99US-0121825.                     |     |
| PR       | 05-MAR-1999;                                                   | 99US-0123180.                     |     |
| PR       | 09-MAR-1999;                                                   | 99US-0123548.                     |     |
| PR       | 23-MAR-1999;                                                   | 99US-0125788.                     |     |
| PR       | 25-MAR-1999;                                                   | 99US-0126264.                     |     |
| PR       | 29-MAR-1999;                                                   | 99US-0126785.                     |     |
| PR       | 01-APR-1999;                                                   | 99US-0127462.                     |     |
| PR       | 06-APR-1999;                                                   | 99US-0128234.                     |     |
| PR       | 08-APR-1999;                                                   | 99US-0128714.                     |     |
| PR       | 16-APR-1999;                                                   | 99US-0129845.                     |     |
| PR       | 19-APR-1999;                                                   | 99US-0130077.                     |     |
| PR       | 21-APR-1999;                                                   | 99US-0130449.                     |     |
| PR       | 23-APR-1999;                                                   | 99US-0130510.                     |     |
| PR       | 28-APR-1999;                                                   | 99US-0130891.                     |     |
| PR       | 30-APR-1999;                                                   | 99US-0131449.                     |     |
| PR       | 30-APR-1999;                                                   | 99US-0132048.                     |     |
| PR       | 30-APR-1999;                                                   | 99US-0132407.                     |     |
| PR       | 04-MAY-1999;                                                   | 99US-0132484.                     |     |
| PR       | 05-MAY-1999;                                                   | 99US-0132485.                     |     |
| PR       | 06-MAY-1999;                                                   | 99US-0132486.                     |     |
| PR       | 06-MAY-1999;                                                   | 99US-0132487.                     |     |
| PR       | 07-MAY-1999;                                                   | 99US-0132863.                     |     |
| PR       | 11-MAY-1999;                                                   | 99US-0134256.                     |     |
| PR       | 14-MAY-1999;                                                   | 99US-0134218.                     |     |
| PR       | 14-MAY-1999;                                                   | 99US-0134219.                     |     |
| PR       | 14-MAY-1999;                                                   | 99US-0134221.                     |     |
| PR       | 14-MAY-1999;                                                   | 99US-0134370.                     |     |
| PR       | 18-MAY-1999;                                                   | 99US-0134768.                     |     |
| PR       | 19-MAY-1999;                                                   | 99US-0134941.                     |     |
| PR       | 20-MAY-1999;                                                   | 99US-0135124.                     |     |
| PR       | 21-MAY-1999;                                                   | 99US-0135353.                     |     |
| PR       | 24-MAY-1999;                                                   | 99US-0135629.                     |     |
| PR       | 25-MAY-1999;                                                   | 99US-0136021.                     |     |
| PR       | 27-MAY-1999;                                                   | 99US-0136392.                     |     |
| PR       | 28-MAY-1999;                                                   | 99US-0136782.                     |     |
| PR       | 01-JUN-1999;                                                   | 99US-0137222.                     |     |
| PR       | 03-JUN-1999;                                                   | 99US-0137548.                     |     |
| PR       | 04-JUN-1999;                                                   | 99US-0137502.                     |     |
| PR       | 07-JUN-1999;                                                   | 99US-0137724.                     |     |
| PR       | 08-JUN-1999;                                                   | 99US-0138094.                     |     |
| PR       | 10-JUN-1999;                                                   | 99US-0138540.                     |     |
| PR       | 10-JUN-1999;                                                   | 99US-0138847.                     |     |
| PR       | 14-JUN-1999;                                                   | 99US-0139119.                     |     |
| PR       | 16-JUN-1999;                                                   | 99US-0139452.                     |     |
| PR       | 16-JUN-1999;                                                   | 99US-0139453.                     |     |
| PR       | 17-JUN-1999;                                                   | 99US-0139492.                     |     |

PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
Pred. No.: 1,08e-21  
Score: 141.00  
Length: 1190  
Matches: 148  
Conservative: 10  
Percent Similarity: 67.23%  
Mismatches: 63  
Best Local Similarity: 62.98%  
Indels: 14  
Query Match: 63.51%  
Gaps: 1  
DB: 21

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x AAC38528 (1-1190)

QY 1 MetSerSerArgTtpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20

Db 59 ATGAGCAGTGGTTCGAGTAGACCGGTACGTGCGAACCTTCTCGCGATATCCGTGAG 118  
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIle 40  
Db 119 AGAGAGGTCGAAGATTGTTTCAGTAAGTATGACCTGTTCTCAATGATTGAAGTT 178  
Qy 41 ProProArgProProGlyTyrAlaPheValGluAspProArgAspAlaAsp 60  
Db 179 CCTCCAAAGGCTCCTGTTATGCAATTCGTTGAGTTTATGATGCTCGGATGCTGAAGAT 238  
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80  
Db 239 GCTATTGATGCTGCTGATGCTATGACTTTGATGGCATCGTTGAGGGTGAATTGGCG 298  
Qy 81 HisGlyGlyArg-----Arg\*\*\* 86  
Db 299 CATGGTGGAGGCGTTCATCAGATGATACCTCGGGTAGTTTCAATGGTGGTGGCCGTG 358  
Qy 87 \*\*\*\*\*SerTyrSerAlaSerArgAlaProSerArgArg 106  
Db 359 GGTGGTGGCGCGCTGCTGATGCGCGGTGATGCTGATGCTGCGGCATCTAGGAGA 418  
Qy 107 SerAspTyrArgValIleuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126  
Db 419 TCAGAGTTTCGTGTTCTAGTCACAGGCTTCCTTCATCTCTCTTGGCAAGATCTCAAG 478  
Qy 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146  
Db 479 GATCACAATGTTAAGAGAGCGGATGCTGTTTCCGAAGTGACCGTATGCTAGAGGG 538  
Qy 147 MetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeu 166  
Db 539 ACAACTGGAGTGTGTTGATTACACCTGCTATGAGACATGAAGTATGCGCTGAAAAGCTC 598  
Qy 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186  
Db 599 GACGACACAGAGTTTCGAAATGCGTTTTCGAATGATATGTCGGGTAGAGAATATGAT 658  
Qy 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer 206  
Db 659 TCAAGGAAGGATTTAGAGTCTTACCGGGGAGATCCTATTCTAAGAGCCGACGCCGC 718  
Qy 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221  
Db 719 AGCCGTGGAGACCGTGCACCCGACAGGACGAGGACGAGGAGC 763

RESULT 8  
AAC37026  
AAC37026 standard; DNA; 1241 BP.  
AAC37026;  
17-OCT-2000 (first entry)  
Arabidopsis thaliana DNA fragment SEQ ID NO: 15916.  
Hybridisation assay; genetic mapping; gene expression control;  
protein identification; signal transduction pathway;  
metabolic pathway; promoter; termination sequence; ss.  
Arabidopsis thaliana.  
EP1033405-A2.  
06-SEP-2000.  
25-FEB-2000; 2000EP-0301439.  
25-FEB-1999; 99US-0121825.  
05-MAR-1999; 99US-0123180.  
09-MAR-1999; 99US-0123548.  
23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140354.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.



[illegible]



```
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149928;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149928;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149928;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149928;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;
```

## Alignment Scores:

Pred. No.: 5,2e-16 Length: 885  
Score: 116.50 Matches: 149  
Percent Similarity: 67.66% Conservative: 10  
Best Local Similarity: 63.40% Mismatches: 62  
Query Match: 52.48% Indels: 15  
DB: 21 Gaps: 2

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x AAC42654 (1-885)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20  
DB 1 ATGAGCAGCGGTCGAGTAGAAGCATTTACGTCGGGAACCTTCGCGGCGATATCCGTGAA 60  
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40  
DB 61 AGAGAAGTTGAAGACTTGTTCAAGTAGACCTGTTTCAATCGATTGAAGATT 120  
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60  
DB 121 CGCGCAGGCGCTCCAGGCTATGCTCGAGTTTGAGATGCTCGTATGCTGATGAT 180  
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80  
DB 181 GCRAITTTATGCGCGTGATGTTATGACITTTGATGGCATCATTTACGGGTGGAAC TAGCT 240  
QY 81 HisGlyGlyArgArg-----\*\*\*\*\* 89  
DB 241 CATGTTGGAGGCGTTTCATCATGATGATCAGCGGTAGTTATAGTGTGCTGCTGCTGCG 300  
QY 90 \*\*\*\*\*SerTyrSerAlaSerAlaSerAlaProSerArgSerArgSerArgSerArg 109  
DB 301 GGTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360  
QY 110 ArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMet 129  
DB 361 CGCGTTGATGTCAGGTTTGCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 130 ArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGly 149  
DB 421 CGTAAAGGAGGAGAGAGTTGTTTCTCAAGTGTTCGTGATGCTGATGATGATGATGATGATG 480  
QY 150 ValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle-----ArgLysLeu 166  
DB 481 ATTGTAGATTATACCAGCTACGAGGACATGAAATATGCGGT-GAGATATAAAAAAGCTC 539  
QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186  
DB 540 GATGACACAGAGTTTCGGAATCGTTTCTCATGAATATGTTCCGGTTAGAGAAATATGAT 599  
QY 187 SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer 206  
DB 600 TCAAGGAGGAGATTCGAGGAGGCCCGCGTGAAGATCCTATTCTAAGAGCGCGAGCCGT 659  
QY 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221  
DB 660 GGACGGAGTCCCGCGGTAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 704

## RESULT 10

AAC35763

ID AAC35763 standard; DNA; 479 BP.

XX

AC AAC35763;

XX

DT 17-OCT-2000 (first entry)

XX

DE Zea mays DNA fragment SEQ ID NO: 11334.

XX

KW Hybridisation assay; Genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic;

KW pathway; promoter; termination sequence; corn; ss.

XX

OS Zea mays subsp. mays.

|    |                              |    |              |               |
|----|------------------------------|----|--------------|---------------|
| XX | XX                           | PR | 01-JUL-1999; | 99US-0141842. |
| PN | EP1033405-A2.                | PR | 01-JUL-1999; | 99US-0142154. |
| XX |                              | PR | 02-JUL-1999; | 99US-0142055. |
| PD | 06-SEP-2000.                 | PR | 06-JUL-1999; | 99US-0142390. |
| XX |                              | PR | 08-JUL-1999; | 99US-0142803. |
| PF | 25-FEB-2000; 2000EP-0301439. | PR | 09-JUL-1999; | 99US-0142920. |
| XX |                              | PR | 12-JUL-1999; | 99US-0142977. |
| XX |                              | PR | 13-JUL-1999; | 99US-0143542. |
| PR | 25-FEB-1999;                 | PR | 14-JUL-1999; | 99US-0143624. |
| PR | 05-MAR-1999;                 | PR | 15-JUL-1999; | 99US-0144005. |
| PR | 09-MAR-1999;                 | PR | 16-JUL-1999; | 99US-0144085. |
| PR | 23-MAR-1999;                 | PR | 16-JUL-1999; | 99US-0144086. |
| PR | 25-MAR-1999;                 | PR | 19-JUL-1999; | 99US-0144325. |
| PR | 29-MAR-1999;                 | PR | 19-JUL-1999; | 99US-0144331. |
| PR | 01-APR-1999;                 | PR | 19-JUL-1999; | 99US-0144332. |
| PR | 06-APR-1999;                 | PR | 19-JUL-1999; | 99US-0144333. |
| PR | 08-APR-1999;                 | PR | 19-JUL-1999; | 99US-0144334. |
| PR | 16-APR-1999;                 | PR | 19-JUL-1999; | 99US-0144335. |
| PR | 19-APR-1999;                 | PR | 20-JUL-1999; | 99US-0144352. |
| PR | 21-APR-1999;                 | PR | 20-JUL-1999; | 99US-0144632. |
| PR | 23-APR-1999;                 | PR | 20-JUL-1999; | 99US-0144884. |
| PR | 23-APR-1999;                 | PR | 20-JUL-1999; | 99US-0144884. |
| PR | 28-APR-1999;                 | PR | 21-JUL-1999; | 99US-0144814. |
| PR | 30-APR-1999;                 | PR | 21-JUL-1999; | 99US-0145086. |
| PR | 30-APR-1999;                 | PR | 21-JUL-1999; | 99US-0145088. |
| PR | 30-APR-1999;                 | PR | 21-JUL-1999; | 99US-0145088. |
| PR | 04-MAY-1999;                 | PR | 22-JUL-1999; | 99US-0145085. |
| PR | 05-MAY-1999;                 | PR | 22-JUL-1999; | 99US-0145087. |
| PR | 06-MAY-1999;                 | PR | 22-JUL-1999; | 99US-0145089. |
| PR | 06-MAY-1999;                 | PR | 22-JUL-1999; | 99US-0145192. |
| PR | 07-MAY-1999;                 | PR | 23-JUL-1999; | 99US-0145145. |
| PR | 11-MAY-1999;                 | PR | 23-JUL-1999; | 99US-0145218. |
| PR | 14-MAY-1999;                 | PR | 23-JUL-1999; | 99US-0145224. |
| PR | 14-MAY-1999;                 | PR | 26-JUL-1999; | 99US-0145276. |
| PR | 14-MAY-1999;                 | PR | 27-JUL-1999; | 99US-0145913. |
| PR | 14-MAY-1999;                 | PR | 27-JUL-1999; | 99US-0145918. |
| PR | 18-MAY-1999;                 | PR | 27-JUL-1999; | 99US-0145919. |
| PR | 18-MAY-1999;                 | PR | 28-JUL-1999; | 99US-0145951. |
| PR | 19-MAY-1999;                 | PR | 02-AUG-1999; | 99US-0146386. |
| PR | 20-MAY-1999;                 | PR | 02-AUG-1999; | 99US-0146388. |
| PR | 21-MAY-1999;                 | PR | 02-AUG-1999; | 99US-0146389. |
| PR | 24-MAY-1999;                 | PR | 03-AUG-1999; | 99US-0147038. |
| PR | 25-MAY-1999;                 | PR | 04-AUG-1999; | 99US-0147204. |
| PR | 27-MAY-1999;                 | PR | 04-AUG-1999; | 99US-0147302. |
| PR | 28-MAY-1999;                 | PR | 05-AUG-1999; | 99US-0147192. |
| PR | 01-JUN-1999;                 | PR | 05-AUG-1999; | 99US-0147260. |
| PR | 03-JUN-1999;                 | PR | 06-AUG-1999; | 99US-0147303. |
| PR | 04-JUN-1999;                 | PR | 06-AUG-1999; | 99US-0147416. |
| PR | 07-JUN-1999;                 | PR | 09-AUG-1999; | 99US-0147493. |
| PR | 08-JUN-1999;                 | PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-JUN-1999;                 | PR | 10-AUG-1999; | 99US-0148171. |
| PR | 10-JUN-1999;                 | PR | 11-AUG-1999; | 99US-0148319. |
| PR | 14-JUN-1999;                 | PR | 12-AUG-1999; | 99US-0148341. |
| PR | 16-JUN-1999;                 | PR | 13-AUG-1999; | 99US-0148565. |
| PR | 16-JUN-1999;                 | PR | 13-AUG-1999; | 99US-0148684. |
| PR | 17-JUN-1999;                 | PR | 16-AUG-1999; | 99US-0149368. |
| PR | 18-JUN-1999;                 | PR | 17-AUG-1999; | 99US-0149175. |
| PR | 18-JUN-1999;                 | PR | 18-AUG-1999; | 99US-0149426. |
| PR | 18-JUN-1999;                 | PR | 20-AUG-1999; | 99US-0149722. |
| PR | 18-JUN-1999;                 | PR | 20-AUG-1999; | 99US-0149723. |
| PR | 18-JUN-1999;                 | PR | 20-AUG-1999; | 99US-0149929. |
| PR | 18-JUN-1999;                 | PR | 23-AUG-1999; | 99US-0149902. |
| PR | 18-JUN-1999;                 | PR | 23-AUG-1999; | 99US-0149930. |
| PR | 18-JUN-1999;                 | PR | 25-AUG-1999; | 99US-0150566. |
| PR | 18-JUN-1999;                 | PR | 26-AUG-1999; | 99US-0150884. |
| PR | 18-JUN-1999;                 | PR | 27-AUG-1999; | 99US-0151065. |
| PR | 18-JUN-1999;                 | PR | 27-AUG-1999; | 99US-0151066. |
| PR | 21-JUN-1999;                 | PR | 27-AUG-1999; | 99US-0151080. |
| PR | 22-JUN-1999;                 | PR | 30-AUG-1999; | 99US-0151303. |
| PR | 22-JUN-1999;                 | PR | 31-AUG-1999; | 99US-0151438. |
| PR | 23-JUN-1999;                 | PR | 01-SEP-1999; | 99US-0151930. |
| PR | 23-JUN-1999;                 | PR | 07-SEP-1999; | 99US-0152363. |
| PR | 24-JUN-1999;                 | PR | 10-SEP-1999; | 99US-0153070. |
| PR | 28-JUN-1999;                 | PR | 13-SEP-1999; | 99US-0153758. |
| PR | 29-JUN-1999;                 | PR | 15-SEP-1999; | 99US-0154018. |
| PR | 30-JUN-1999;                 | PR |              |               |

PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162143.

Alignment Scores:  
Pred. No.: 2,57e-12 Length: 479  
Score: 98.00 Matches: 88  
Percent Similarity: 79.03% Conservative: 10  
Best Local Similarity: 70.97% Mismatches: 26  
Query Match: 44.14% Indels: 0  
DB: 21 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x AAC35763 (1-479)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20  
Db 108 ATGACCCGGCGGAACGGGTGTACGATCTACGTGGCAACCTCCCGCGGACATCCGGAG 167  
QY 21 CysGluValGluAspLeuPheTyrIleValAspIleAspLeuIle 40  
Db 168 AGGGAAGTGGATGATCTCTTCAAGATGCGCATAGTGGAAATGACTTGAAAT 227  
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60  
Db 228 CCACCAAGCCCTCCGGTTTGTCTTTTGTGAGTTTGGAGCGGACGATGCTGAAGAT 287  
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80  
Db 288 GCAATATATGCGGTGATGATACAACTTTTGTGCGCATAGTTTGGGGTGAATTAGCC 347  
QY 81 HisGlyGlyArgArg\*\*\*\*\*SerTyrSerAlaSer 100  
Db 348 CATGTGACGAGGCACATCTTTTGTATCGATCTAGCAGCTATAGCAGTGTGGACAA 407  
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAla 120

Db 408 CCGGTGGCTCAAAACGTTGATTACCGTGTATGGTTACTGGATTACCTTCTTCAGCA 467  
QY 121 SerTrpGlnAsp 124  
Db 468 TCATGGCAAGAT 479

RESULT 11  
AAH02904  
ID AAH02904 standard; DNA; 1428 BP.  
XX AAH02904;  
XX 15-JUN-2001 (first entry)  
XX Human shear stress-response coding sequence SEQ ID NO: 61.  
XX Human; shear stress-response protein; vascular disease;  
KW arteriosclerosis; ds.  
XX Homo sapiens.  
OS  
PN WO200125427-A1.  
XX  
PD 12-APR-2001.  
XX  
PF 02-OCT-2000; 2000WO-JP06840.  
XX  
PR 01-OCT-1999; 99JP-0280976.  
XX  
PA (KYOW) KYOWA HAKKO KOGYO KK.  
PA (NOJI/) NOJIMA H.  
XX  
PI Nojima H, Yoshiue H, Obayashi M, Ota T, Kawabata A, Sakurada K;  
PI Kuga T, Sekine S, Nakamura Y, Sugano S;  
XX  
XX WPI; 2001-366308/27.  
DR P-PSDB; AAB90781.  
XX  
XX DNA sequences, proteins encoded by them and antibodies against them  
PT useful in diagnosis and treatment of vascular disease caused by  
PT arteriosclerosis -  
XX  
PS Claim 20; Page 386-388; 678pp; Japanese.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human shear stress response proteins. These are useful in the  
CC diagnosis, treatment and screening of vascular diseases caused by  
CC arteriosclerosis, including heart failure, post-PTCA restenosis and  
CC hypertension.  
XX  
SQ Sequence 1428 BP; 341 A; 296 C; 400 G; 391 T; 0 other;

Alignment Scores:  
Pred. No.: 6,7e-10 Length: 1428  
Score: 95.00 Matches: 108  
Percent Similarity: 56.46% Conservative: 10  
Best Local Similarity: 51.67% Mismatches: 85  
Query Match: 42.79% Indels: 6  
DB: 22 Gaps: 2

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x AAH02904 (1-1428)

QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28  
Db 176 ATCTAGTGGTAACTTACCTCCAGACATCCGAACCAAGGACATTGAGACGCTGTCTAC 235  
QY 29 LysTyrGlyProIleValAspIleAspLeuIleProProArgProGlyTyrAla 48  
Db 236 AAATACGGCGCTATCCGGACATCCACCTCAAGAAATCCCGCGGGGACCGCCCTTCGCC 295  
QY 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68

|           |                                                                 |                                                              |     |    |              |               |
|-----------|-----------------------------------------------------------------|--------------------------------------------------------------|-----|----|--------------|---------------|
| Db        | 296                                                             | TTGGTTGAGTTCGAGGACCCGCGAGACCGCGGTGATGTCGCGACGGCTAT           | 355 | PR | 04-MAY-1999; | 99US-0132484. |
| Qy        | 69                                                              | AppPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg     | 85  | PR | 05-MAY-1999; | 99US-0132485. |
| Db        | 356                                                             | GATTACGATGGTACCGTCTGCGGGTGGAGTTTCCTCGAAGCGGCGGTGGAACAGCCGA   | 415 | PR | 06-MAY-1999; | 99US-0132486. |
| Qy        | 86                                                              | *****SerTyrSerAlaSerArgAlaProSerArg                          | 105 | PR | 07-MAY-1999; | 99US-0132487. |
| Db        | 416                                                             | GGCGGGCGGGGTGGAGTGGCGGAGTCCCGGAGTGGTATGGCCCCCATCCAGG         | 475 | PR | 11-MAY-1999; | 99US-0132488. |
| Qy        | 106                                                             | ArgSerAspTyrArgValLeuValThrGlyLeuProPheSerAlaSerTyrGlnAspLeu | 125 | PR | 14-MAY-1999; | 99US-0132489. |
| Db        | 476                                                             | CGGTCTGAAACAGAGTGGTGTCTGCTGACTGCTCCAGTGGAGTGGCAGGATTTA       | 535 | PR | 14-MAY-1999; | 99US-0132490. |
| Qy        | 126                                                             | LysAspHisMetArgLysValGlyAspValCysPheSerGluValPheProAspArgLys | 145 | PR | 18-MAY-1999; | 99US-0132491. |
| Db        | 536                                                             | AAGGATCATCGTGCAGCAGTGTATGTTATGCTGATGTTTACCGAGAT-----         | 589 | PR | 18-MAY-1999; | 99US-0132492. |
| Qy        | 146                                                             | GlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLys    | 165 | PR | 20-MAY-1999; | 99US-0132493. |
| Db        | 590                                                             | ---GGCACCTGGTGTCTGAGTTTGTACGGAAAGAGATATGACCTATGCAATTCGAAA    | 646 | PR | 21-MAY-1999; | 99US-0132494. |
| Qy        | 166                                                             | LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr | 185 | PR | 24-MAY-1999; | 99US-0132495. |
| Db        | 647                                                             | CTGGATAACACTAAGTTTAGATCTCATGAGGAGAACTGCCTACATCGGGTTAAAGTT    | 706 | PR | 25-MAY-1999; | 99US-0132496. |
| Qy        | 186                                                             | GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg | 205 | PR | 27-MAY-1999; | 99US-0132497. |
| Db        | 707                                                             | GATGGCCCGAGGTCAGTTCAGTATGGAAGATCTCGATCTCGAAGCGGTAGTGTAGCAGA  | 766 | PR | 28-MAY-1999; | 99US-0132498. |
| Qy        | 206                                                             | SerArgGlyProSerCysSerTyrSer                                  | 214 | PR | 01-JUN-1999; | 99US-0132499. |
| Db        | 767                                                             | AGCCGTAGCAGAACCAACAGCAGGAGT                                  | 793 | PR | 03-JUN-1999; | 99US-0132500. |
| RESULT 12 |                                                                 |                                                              |     |    |              | 99US-0132501. |
| AAC44527  |                                                                 |                                                              |     |    |              | 99US-0132502. |
| ID        | AAC44527 standard; DNA; 498 BP.                                 |                                                              |     |    |              | 99US-0132503. |
| XX        | AAC44527;                                                       |                                                              |     |    |              | 99US-0132504. |
| AC        | AAC44527;                                                       |                                                              |     |    |              | 99US-0132505. |
| XX        | AAC44527;                                                       |                                                              |     |    |              | 99US-0132506. |
| DT        | 18-OCT-2000 (first entry)                                       |                                                              |     |    |              | 99US-0132507. |
| DE        | Zea mays DNA fragment SEQ ID NO: 43156.                         |                                                              |     |    |              | 99US-0132508. |
| XX        | Zea mays DNA fragment SEQ ID NO: 43156.                         |                                                              |     |    |              | 99US-0132509. |
| XX        | Hybridisation assay; genetic mapping; gene expression control;  |                                                              |     |    |              | 99US-0132510. |
| KW        | protein identification; signal transduction pathway; metabolic; |                                                              |     |    |              | 99US-0132511. |
| XX        | pathway; promoter; termination sequence; corn; ss.              |                                                              |     |    |              | 99US-0132512. |
| OS        | Zea mays subsp. mays.                                           |                                                              |     |    |              | 99US-0132513. |
| XX        | EP1033405-A2.                                                   |                                                              |     |    |              | 99US-0132514. |
| PN        | EP1033405-A2.                                                   |                                                              |     |    |              | 99US-0132515. |
| XX        | EP1033405-A2.                                                   |                                                              |     |    |              | 99US-0132516. |
| PD        | 06-SEP-2000.                                                    |                                                              |     |    |              | 99US-0132517. |
| XX        | 25-FEB-2000; 2000EP-0301439.                                    |                                                              |     |    |              | 99US-0132518. |
| XX        | 25-FEB-1999; 99US-0121825.                                      |                                                              |     |    |              | 99US-0132519. |
| PR        | 05-MAR-1999; 99US-0123180.                                      |                                                              |     |    |              | 99US-0132520. |
| PR        | 09-MAR-1999; 99US-0123548.                                      |                                                              |     |    |              | 99US-0132521. |
| PR        | 23-MAR-1999; 99US-0125788.                                      |                                                              |     |    |              | 99US-0132522. |
| PR        | 25-MAR-1999; 99US-0126264.                                      |                                                              |     |    |              | 99US-0132523. |
| PR        | 29-MAR-1999; 99US-0126785.                                      |                                                              |     |    |              | 99US-0132524. |
| PR        | 01-APR-1999; 99US-0127462.                                      |                                                              |     |    |              | 99US-0132525. |
| PR        | 06-APR-1999; 99US-0128234.                                      |                                                              |     |    |              | 99US-0132526. |
| PR        | 16-APR-1999; 99US-0128714.                                      |                                                              |     |    |              | 99US-0132527. |
| PR        | 19-APR-1999; 99US-0129845.                                      |                                                              |     |    |              | 99US-0132528. |
| PR        | 21-APR-1999; 99US-0130077.                                      |                                                              |     |    |              | 99US-0132529. |
| PR        | 23-APR-1999; 99US-0130449.                                      |                                                              |     |    |              | 99US-0132530. |
| PR        | 28-APR-1999; 99US-0130510.                                      |                                                              |     |    |              | 99US-0132531. |
| PR        | 30-APR-1999; 99US-0130891.                                      |                                                              |     |    |              | 99US-0132532. |
| PR        | 30-APR-1999; 99US-0131449.                                      |                                                              |     |    |              | 99US-0132533. |
| PR        | 30-APR-1999; 99US-0132048.                                      |                                                              |     |    |              | 99US-0132534. |
| PR        | 30-APR-1999; 99US-0132407.                                      |                                                              |     |    |              | 99US-0132535. |



XX PD 11-APR-2002.  
XX PF 02-OCT-2001; 2001WO-US30589.  
XX PR 02-OCT-2000; 2000US-237054P.  
XX PA (GENE-) GENE LOGIC INC.  
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX DR WPI; 2002-426119/45.  
XX PT Diagnosing and detecting the progression of liver cancer,  
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
XX PT involves detecting the level of expression of two or more genes in a  
XX PT liver tissue sample  
XX PS Claim 1; SEQ ID NO 3343; 298pp; English.  
XX CC The invention relates to a novel method for diagnosing and detecting the  
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
XX CC tumor in a patient, and differentiating metastatic liver cancer from  
XX CC hepatocellular carcinoma in a patient, involving detecting the level of  
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a  
XX CC tissue sample. The method of the invention has hepatotropic, and  
XX CC cytostatic activity. The method is useful for diagnosing and detecting  
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
XX CC liver carcinoma in a patient. The method is useful for identifying  
XX CC expression profiles which serve as useful diagnostic markers as well as  
XX CC markers that can be used to monitor disease states, disease progression,  
XX CC drug toxicity, drug efficacy and drug metabolism.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 1069 BP; 245 A; 229 C; 310 G; 285 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 1-83e-08 Length: 1069  
Score: 87.50 Matches: 90  
Percent Similarity: 48.54% Conservative: 10  
Best Local Similarity: 43.69% Mismatches: 101  
Query Match: 39.41% Indels: 5  
DB: 24 Gaps: 1  
  
US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x ABN96845 (1-1069)  
QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28  
DB 98 AACTAGTGGGAACTTCGACCGACGTGCGGAGAGGACTTGGAGGACCTGTCTTAC 157  
QY 29 LysTyrGlyProIleValAspLeuAspLeuLysIleProProArgProProGlyTyrAla 48  
DB 158 AAGTACGGCGGATCGCGAGATCGAGCTCAAGAACCGGACGGCTCGTGGCCCTCGCC 217  
QY 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68  
DB 218 TTCGTGCGTTCGAGACCCCGGAGATCGAGAGTGTCTATTATCGAAGAAATCGTAT 277  
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg\*\*\*\*\* 88  
DB 278 GATATGGCCAGTTCGGCTTCGTGTGGAG-----TTCCCGAGGACTTAT 322  
QY 89 \*\*\*\*\*SertySerAlaSerArgAlaProSerArgArgSerAsp 108  
DB 323 GGAGGTGGGGTGGTGGCCCGTGTGGGAGGAATGGCCCTCTTCAAGAAGATCTGAT 382  
QY 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis 128  
DB 383 TTCGAGGTTCTTGTTCAGGACTTCCTCGTCAGGCAGCTGGCAGGACTGAAGGATCAC 442  
QY 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 149

DB 443 ATGGGAGAGCTGGGATGCTCTGTATGCTGATGTCAGAGAGATGGAGTGGGATGGTC 502  
QY 149 GlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeuAspAla 168  
DB 503 GAGTATCTCAGAAAAGAGACATGGAATATGCCCTGCGTAAACTGGATGACACCAATTC 562  
QY 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188  
DB 563 CGTCTCTCATGAGGTTGAAACTTCTTATCCGAGTTTATCTCTGAGAGAACACACGCTAT 622  
QY 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208  
DB 623 GGCTACTCAGGTTCTGGTCTGGTCAAGGGCGGTGACTCTCATACCAAGACAGGGGT 682  
QY 209 ProSerCysSerTyrSer 214  
DB 683 TCCCCACACTACTTCTCT 700  
  
RESULT 14  
AAF16226  
ID AAF16226 standard; cDNA; 1162 BP.  
XX AC AAF16226;  
XX DT 13-MAR-2001 (first entry)  
XX CC Human prostate cancer antigen nucleotide sequence SEQ ID NO:661.  
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
XX KW neuroprotective; cytostatic; cardiocactive; immunomodulatory; muscular;  
XX KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
XX KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
XX KW wound; infectious disease; ss.  
XX OS Homo sapiens.  
XX FN WO200055174-A1.  
XX PD 21-SEP-2000.  
XX PF 08-MAR-2000; 2000WO-US05988.  
XX PR 12-MAR-1999; 99US-0124270.  
XX FA (HUMA-) HUMAN GENOME SCI INC.  
XX FA (ROSE/) ROSEN C A.  
XX PI Rosen CA, Ruben SM;  
XX DR WPI; 2000-587513/55.  
XX DR P-PSDB; AAB57023.  
XX PT Prostate cancer associated gene sequences, referred to as prostate  
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of  
XX PT disorders such as prostate cancer -  
XX PS Claim 1; Page 1098-1099; 2338pp; English.  
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated  
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,  
XX CC cardiocactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
XX CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
XX CC and can be used in gene therapy. The prostate cancer antigen  
XX CC polynucleotides may be used for detection of prostate cancer, chromosome  
XX CC identification, as chromosome markers, and for numerous other diagnostic  
XX CC or research purposes. The prostate cancer antigens may be used to treat  
XX CC disorders such as neural, immune, muscular, reproductive,  
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
XX CC AAB57303 represent sequences used in the exemplification of the present



```
CC invention.
XX
SQ Sequence 1162 BP; 280 A; 252 C; 340 G; 288 T; 2 other;

Alignment Scores:
Pred. No.: 2,45e-08 Length: 1162
Score: 87.50 Matches: 90
Percent Similarity: 48.54% Conservatives: 10
Best Local Similarity: 43.69% Mismatches: 101
Query Match: 39.41% Indels: 5
DB: 21 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x AAF16226 (1-1162)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 143 ATCTACGTGGGAACCTCCGACCGACGTGGCGGAGAGGACTTGGAGGACCTGTCTAC 202
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 203 AAGTACGCGCCATCCCGAGATGAGTCAAGAACCGGACCGCCCTCGTCCCTTCGCC 262
Qy 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
Db 263 TTCGTGGCTTCGAGGACCCCGAGATCGAGAGTCTATTATGGAAGAAATGGTTAT 322
Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 88
Db 323 GATTATGCCAGTGTGGCTTCGTGTGGAG-----TTCGCCAGGACTTAT 367
Qy 89 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAsp 108
Db 368 GGAGTCCGGGTGGTGGCCCGTGGTGGGAGGAATGGCGCTCTACAGAAGATCTGAT 427
Qy 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis 128
Db 428 TTCGAGTTCGTGTTCCAGGACTTCCTCCGFCAGGACGTGGCAGGACCTGAAGATCAC 487
Qy 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 148
Db 488 ATGGAGAGCTGGGATGCTGTATGCTGATGTGCAGAGAGTGGAGTGGGATGGTC 547
Qy 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla 168
Db 548 GAGTATCTCAGAAAAGAGACATGGAATATGCCCTGGCTAACTGGATGACACAAATTC 607
Qy 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188
Db 608 CGCTCTCATGAGGTGAAACTTCATATCCGAGCTTTATCCTGAGAGAAGCACCAGCTAT 667
Qy 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208
Db 668 GGCTACTCACGGTCTCGGTCTGGGTCAAGGGCGCGTGACTCTCCATACCAAGCAGGGT 727
Qy 209 ProSerCysSerTyrSer 214
Db 728 TCCCCACACTACTTCTCT 745

RESULT 15
ABV30279
ID ABV30279 standard; cDNA; 1218 BP.
XX
XX AC ABV30279;
XX
XX DT 16-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 30270.
XX
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX OS Homo sapiens.
XX
```

```
PN WO200150860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PF 17-FEB-2000; 2000US-183119P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 6574-6575; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 1218 BP; 280 A; 275 C; 357 G; 305 T; 1 other;
SQ

Alignment Scores:
Pred. No.: 2,89e-08 Length: 1218
Score: 87.50 Matches: 90
Percent Similarity: 48.54% Conservatives: 10
Best Local Similarity: 43.69% Mismatches: 101
Query Match: 39.41% Indels: 5
DB: 23 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x ABV30279 (1-1218)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 180 ATCTACGTGGGAACCTCCGACCGACGTGGCGGAGAGGACTTGGAGGACCTGTCTAC 239
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 240 AAGTACGCGCCATCCCGAGATCGAGTCAAGAACCGGACCGGCTCGTCCCTTCGCC 299
Qy 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
Db 300 TTCGTGGCTTCGAGGACCCCGAGATGCGAGGATGCTATTATGGAAGAAATGGTTAT 359
Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArgArg***** 88
Db 360 GATTATGCCAGTGTGGCTTCGTGTGGAG-----TTCGCCAGGACTTAT 404
Qy 89 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAsp 108
Db 405 GGAGTCCGGGTGGTGGCCCGCTGGTGGGAGGAATGGGCTCTCTACAGAAGATCTGAT 464
```

Qy 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAspHis 128  
Db 465 TTCCGAGTTCTTGTTCAGGACTTCCTCCGTCAGGCAGCTGGCAGGACCTGAGAGATCAC 524  
Qy 129 MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer 148  
Db 525 ATGCGAGAAGCTGGGGATGCTGTATGCTGATGTCAGAGAAGGATGGAGTGGGGATGGTC 584  
Qy 149 GlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeuAspAla 168  
Db 585 GAGTATCTCAGAAAGACATGGAATATGCCCTCGTTAAACTGGATGACACCAAAATTC 644  
Qy 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188  
Db 645 CGCTCTCATGAGGTGAAACTTCTACATCGAGTTTATCCTGAGAGAAGCACCAGCTAT 704  
Qy 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly 208  
Db 705 GGCTACTCACGGTCTGGTCTGGGTCAAGGGCGGTGACTCTCCATACCAAGCAGGGGT 764  
Qy 209 ProSerCysSerTyrSer 214  
Db 765 TCCCCACACTACTTCTCT 782

Search completed: February 4, 2004, 21:26:36  
Job time : 271 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 20:09:39 ; Search time 349 Seconds  
(without alignment)  
2343.170 Million cell updates/sec

Title: US-10-014-927-19MOD\_COPY\_1\_222  
Perfect score: 222  
Sequence: 1 MSSRWNRRTYVGNLPDIRK.....RSRSGPGSCSYSSKRSVSP 222

Scoring table: UNITARY2  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues 4899406

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/BAUM927/runat\_04022004\_131334\_2737/app\_query.fasta\_1.391  
-DB=Published Applications\_NA -QFWT=fastap -SUFFIX=nrpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=unitary2  
-TRANS=human40.cai -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=BAUM927 @CGN 1.1 221 @runat\_04022004\_131334\_2737  
-NCFU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/prodata/1/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/1/pubna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/1/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/prodata/1/pubna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/prodata/1/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/prodata/1/pubna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/prodata/1/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/prodata/1/pubna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/1/pubna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/prodata/1/pubna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/1/pubna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/prodata/1/pubna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/prodata/1/pubna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/prodata/1/pubna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/prodata/1/pubna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/prodata/1/pubna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/prodata/1/pubna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/prodata/1/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Score | Query Match Length | ID | Description |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|--------------------|----|-------------|
| -----                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |       |                    |    |             |
| *<br>US-09-938-842A-1199<br>; Sequence 1199, Application US/09938842A<br>; Patent No. US20020160378A1<br>; GENERAL INFORMATION:<br>; APPLICANT: Harper, Jeff<br>; APPLICANT: Kreps, Joel<br>; APPLICANT: Wang, Xun<br>; APPLICANT: Zhu, Tong<br>; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING<br>; FILE REFERENCE: SCRIPI300-3<br>; CURRENT APPLICATION NUMBER: US/09/938,842A<br>; CURRENT FILING DATE: 2001-08-24<br>; PRIOR APPLICATION NUMBER: US 60/227,866<br>; PRIOR FILING DATE: 2000-08-24<br>; PRIOR APPLICATION NUMBER: US 60/264,647<br>; PRIOR FILING DATE: 2001-01-16<br>; PRIOR APPLICATION NUMBER: US 60/300,111<br>; PRIOR FILING DATE: 2001-06-22<br>; NUMBER OF SEQ ID NOS: 5379<br>; SEQ ID NO 1199<br>; LENGTH: 762<br>; TYPE: DNA |       |                    |    |             |

|    |      |      |         |    |                     |                   |
|----|------|------|---------|----|---------------------|-------------------|
| 1  | 207  | 93.2 | 762     | 10 | US-09-938-842A-1199 | Sequence 1199, Ap |
| 2  | 207  | 93.2 | 762     | 12 | US-09-938-842A-1199 | Sequence 1199, Ap |
| 3  | 95   | 42.8 | 3299    | 13 | US-10-006-285-513   | Sequence 513, App |
| 4  | 87.5 | 39.4 | 1089    | 10 | US-09-880-107-3342  | Sequence 3342, Ap |
| 5  | 87.5 | 39.4 | 1162    | 10 | US-09-925-300-661   | Sequence 661, App |
| 6  | 87.5 | 39.4 | 1475    | 12 | US-10-264-049-273   | Sequence 273, App |
| 7  | 87.5 | 39.4 | 1717    | 15 | US-10-171-581-9     | Sequence 9, Appli |
| 8  | 81   | 36.5 | 345     | 9  | US-09-770-791-783   | Sequence 783, App |
| 9  | 73.5 | 33.1 | 446     | 15 | US-10-060-036-52    | Sequence 52, Appl |
| 10 | 69   | 31.1 | 759     | 9  | US-09-910-943-651   | Sequence 551, App |
| 11 | 69   | 31.1 | 2601    | 13 | US-10-101-510-334   | Sequence 334, App |
| 12 | 65   | 29.3 | 742     | 9  | US-09-910-943-539   | Sequence 539, App |
| 13 | 62   | 27.9 | 2167    | 12 | US-10-439-703-38    | Sequence 38, Appl |
| 14 | 57   | 25.7 | 1579    | 12 | US-10-388-934-197   | Sequence 197, App |
| 15 | 56   | 25.0 | 4044    | 14 | US-10-014-927-18    | Sequence 18, Appl |
| 16 | 55.5 | 25.2 | 1534    | 14 | US-10-044-090-288   | Sequence 288, App |
| 17 | 55.5 | 25.0 | 3697    | 14 | US-10-002-600-58    | Sequence 58, Appl |
| 18 | 53   | 23.9 | 3026    | 11 | US-09-919-039-314   | Sequence 314, App |
| 19 | 51   | 23.0 | 1750    | 9  | US-09-925-302-316   | Sequence 316, App |
| 20 | 50.5 | 22.7 | 489     | 11 | US-09-918-995-23144 | Sequence 23144, A |
| 21 | 50   | 22.5 | 9027    | 13 | US-09-930-213-304   | Sequence 304, App |
| 22 | 50   | 22.5 | 24387   | 11 | US-09-764-891-9844  | Sequence 9844, Ap |
| 23 | 49   | 22.1 | 2212    | 9  | US-09-925-297-142   | Sequence 142, App |
| 24 | 48   | 21.6 | 307     | 9  | US-09-294-093B-1314 | Sequence 1314, Ap |
| 25 | 48   | 21.6 | 2138    | 12 | US-10-369-493-27564 | Sequence 27664, A |
| 26 | 48   | 21.6 | 9400    | 15 | US-10-270-333-7     | Sequence 7, Appli |
| 27 | 48   | 21.6 | 9810    | 11 | US-09-373-658-35    | Sequence 35, Appl |
| 28 | 48   | 21.6 | 9810    | 12 | US-09-989-687-35    | Sequence 35, Appl |
| 29 | 48   | 21.4 | 9025608 | 15 | US-10-156-761-1     | Sequence 1, Appli |
| 30 | 47.5 | 21.4 | 746     | 9  | US-09-910-943-530   | Sequence 530, App |
| 31 | 47   | 21.2 | 282     | 12 | US-09-864-408A-343  | Sequence 343, App |
| 32 | 47   | 21.2 | 831     | 12 | US-10-369-493-35589 | Sequence 35589, A |
| 33 | 47   | 21.2 | 955     | 13 | US-10-117-723-524   | Sequence 524, App |
| 34 | 47   | 21.2 | 955     | 15 | US-10-037-270-524   | Sequence 524, App |
| 35 | 47   | 21.2 | 1128    | 15 | US-10-156-761-2339  | Sequence 2339, Ap |
| 36 | 47   | 21.2 | 1140    | 15 | US-10-156-761-1433  | Sequence 1433, Ap |
| 37 | 47   | 21.2 | 1191    | 12 | US-10-369-493-34569 | Sequence 34569, A |
| 38 | 47   | 21.2 | 1221    | 13 | US-10-214-446-1     | Sequence 1, Appli |
| 39 | 47   | 21.2 | 1686    | 15 | US-10-156-761-3852  | Sequence 3852, Ap |
| 40 | 47   | 21.2 | 2394    | 9  | US-09-815-242-4016  | Sequence 4016, Ap |
| 41 | 47   | 21.2 | 3573    | 12 | US-10-369-493-42233 | Sequence 42233, A |
| 42 | 47   | 21.2 | 3933    | 12 | US-10-369-493-32532 | Sequence 32532, A |
| 43 | 47   | 21.2 | 4175    | 15 | US-10-307-019-2     | Sequence 2, Appli |
| 44 | 47   | 21.2 | 5007    | 15 | US-10-307-019-3     | Sequence 3, Appli |
| 45 | 47   | 21.2 | 7893    | 14 | US-10-077-130-3     | Sequence 3, Appli |



FILE REFERENCE: PA-0039 US  
 CURRENT APPLICATION NUMBER: US/10/006,285  
 CURRENT FILING DATE: 2001-12-05  
 NUMBER OF SEQ ID NOS: 514  
 SOFTWARE: PERL Program  
 SEQ ID NO 513  
 LENGTH: 3299  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. US20030165854A1 411426.42  
 US-10-006-285-513

Alignment Scores:  
 Pred. No.: 1,59e-11 Length: 3299  
 Score: 95.00 Matches: 108  
 Percent Similarity: 56.46% Conservative: 10  
 Best Local Similarity: 51.67% Mismatches: 85  
 Query Match: 42.79% Indels: 6  
 Gaps: 2

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-10-006-285-513 (1-3299)

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| QY | 9   | IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr | 28  |
| DB | 203 | ATCTAGTGGGTAACTTACCTCCAGACATCCGAACCAAGACATTGAGGACGTCTCTAC    | 262 |
| QY | 29  | LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla | 48  |
| DB | 263 | AAATACGGCGCTATCCGCGACATCGACCTCAAGATCGCGGGGACCGCCCTTCGCC      | 322 |
| QY | 49  | PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr    | 68  |
| DB | 323 | TTCTGTTAGTTCCAGGACCCGCGAGACGCGGAAGACGCGGTATGTCGACGCGTAT      | 382 |
| QY | 69  | AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg             | 85  |
| DB | 383 | GATTACGATGGTACCGTCTCGCGGTGGAGTTTCTCGAAGCGCGCTCGGAACAGCGCA    | 442 |
| QY | 86  | *****SerTyrSerAlaSerArgAlaProSerArg                          | 105 |
| DB | 443 | GGCGGGCGGGGTGGAGTGGCGGAGCTCCCGAGTCCGCTATGGCCCCCATCCAG        | 502 |
| QY | 106 | ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu | 125 |
| DB | 503 | CGGTCTGAAACACAGAGTGGTGTCTCTGGACTCCCTCAAGTGAAGTTGGCAGGATTTA   | 562 |
| QY | 126 | LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys | 145 |
| DB | 563 | AAGATCAGATGCGTGAAGCAGGTGATGATGATGATGATGATGATGATGATGATGATGAT  | 616 |
| QY | 146 | GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys | 165 |
| DB | 617 | ---GGCAGTGGTGGTGGAGTTGTACGGAAGAGATATGACCTATGCGAGTTCAAAA      | 673 |
| QY | 166 | LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr | 185 |
| DB | 674 | CTGGATAACACTAAGTTTAGATCTCATGAGGAGAACTCCCTACATCCGGGTAAAGTT    | 733 |
| QY | 186 | GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg | 205 |
| DB | 734 | GATGGGCCCAAGATCCAAAGTTATGGAAGATCTCGATCTCGAAGCGCTAGTCTGATCAGA | 793 |
| QY | 206 | SerArgGlyProSerCysSerTyrSer                                  | 214 |
| DB | 794 | AGCGTAGCAGACACACAGGAGT                                       | 820 |

RESULT 4  
 US-09-880-107-3342  
 Sequence 3342, Application US/09880107  
 Patent No. US20020142981A1  
 GENERAL INFORMATION:

APPLICANT: Horne, Darci T.  
 APPLICANT: Vockley, Joseph G.  
 APPLICANT: Scherf, Uwe  
 APPLICANT: Gene Logic, Inc.  
 TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 FILE REFERENCE: 44921-5028-WO  
 CURRENT APPLICATION NUMBER: US/09/880,107  
 CURRENT FILING DATE: 2001-06-14  
 PRIOR APPLICATION NUMBER: US 60/211,379  
 PRIOR FILING DATE: 2000-06-14  
 PRIOR APPLICATION NUMBER: US 60/237,054  
 PRIOR FILING DATE: 2000-10-02  
 NUMBER OF SEQ ID NOS: 3950  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3342  
 LENGTH: 1069  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: Genbank Accession No. US20020142981A1 U30825  
 US-09-880-107-3342

Alignment Scores:  
 Pred. No.: 3,23e-11 Length: 1069  
 Score: 87.50 Matches: 90  
 Percent Similarity: 48.54% Conservative: 10  
 Best Local Similarity: 43.69% Mismatches: 101  
 Query Match: 39.41% Indels: 5  
 Gaps: 1

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-880-107-3342 (1-1069)

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| QY | 9   | IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr | 28  |
| DB | 98  | ATCTAGTGGGAACTTCCGACCGAGCTGCGGAGAGGACTTGGAGGACCTGTTCTAC      | 157 |
| QY | 29  | LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla | 48  |
| DB | 158 | AAGTACGGCGCATCCGCGAGATCGAGCTCAAGAACCGGACCGCCCTCGTCCCTTCGCC   | 217 |
| QY | 49  | PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr    | 68  |
| DB | 218 | TTCTGCGCTTCGAGGACCCCGGAGATGCGAGATGCTATTTATGGAAGAAATGGTTAT    | 277 |
| QY | 69  | AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg             | 88  |
| DB | 278 | GATTATGCCAGTGTGGCTTCGTGTGAG-----TTCCCCAGGACTTAT              | 322 |
| QY | 89  | *****SerTyrSerAlaSerArgAlaProSerArgSerArg                    | 108 |
| DB | 323 | GGAGTGGGGTGGTGGCGCCCGTGGGAGGAATGGGCTCTACAGAAGATCTGAT         | 382 |
| QY | 109 | TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis | 128 |
| DB | 383 | TTCCGAGTCTTGTTCAGGACTTCTCCGTCAGGAGCTGGCAGGACCTGAAGGATCAC     | 442 |
| QY | 129 | MetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer | 148 |
| DB | 443 | ATGCGAGAAGTCTGGGATGCTGTTATGCTGATGTCAGAAGATGAGTGGGATGGTGG     | 502 |
| QY | 149 | GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla | 168 |
| DB | 503 | GAGTATCTCAGAAAGAAGACATGGAATATGCCCTCGGTAACCTGGATGACACCAATTC   | 562 |
| QY | 169 | ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg | 188 |
| DB | 563 | CGCTCTCATGAGGGTGAACCTTCTACATCCGAGTTTATCTCTGAGAGAAGCAGCAGTAT  | 622 |
| QY | 189 | SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly | 208 |
| DB | 623 | GGCTACTCAGGTCGCGTCTGGGTCAAGGGGCGGTGACTCTCCATACCAAGAGCGGGT    | 682 |
| QY | 209 | ProSerCysSerTyrSer                                           | 214 |



Db 986 ATGCGAAGAGTGGGATCTCTGTTATGCTGATGTCAGAAAGATGGAGTGGGGATGTC 927  
 QY 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAla 168  
 Db 926 GAGTATCTCAGAAAAGAACACATGGAATATGCCCTGCCGTAAACTGGATCAGACCAATTC 867  
 QY 169 ThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArg 188  
 Db 866 CGCTCTCATGAGGTGAACCTTCCTACATCCAGATTATCTGAGAGAGACACAGCTAT 807  
 QY 189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGly 208  
 Db 806 GGCTACTCAGGTCTCGGCTCGGTCAGGGCCGCTGACTCTCCATACCAAGCAGGGGT 747  
 QY 209 ProSerCysSerTyrSer 214  
 Db 746 TCCACACTACTTCTCT 729

RESULT 7  
 US-10-171-581-9  
 ; Sequence 9, Application US/10171581  
 ; Publication No. US20030104426A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dai, Hongyue  
 ; APPLICANT: Linsley, Peter  
 ; APPLICANT: Mao Mao  
 ; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia  
 ; FILE REFERENCE: 9301-157-999  
 ; CURRENT APPLICATION NUMBER: US/10/171,581  
 ; CURRENT FILING DATE: 2002-06-14  
 ; PRIOR APPLICATION NUMBER: 60/298,914  
 ; PRIOR FILING DATE: 2001-06-18  
 ; NUMBER OF SEQ ID NOS: 366  
 ; SEQ ID NO 9  
 ; LENGTH: 1717  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; PUBLICATION INFORMATION:  
 ; DATABASE ACCESSION NUMBER: M72709  
 ; DATABASE ENTRY DATE: 2001-06-18  
 US-10-171-581-9

Alignment Scores:  
 Pred. No.: 1,96e-10 Length: 1717  
 Score: 87.50 Matches: 89  
 Percent Similarity: 49.75% Conservative: 10  
 Best Local Similarity: 44.72% Mismatches: 97  
 Query Match: 39.41% Indels: 3  
 DB: 15 Gaps: 1

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-10-171-581-9 (1-1717)  
 QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28  
 Db 148 ATCTAGTGGTAACTTACCTCCAGACATCCGAACCAAGACATTCGAGACGTGTTCTAC 207  
 QY 29 LysTyrGlyProIleValAspIleAspLeuIleProProArgProProGlyTyrAla 48  
 Db 208 AAATACGGCGCTATCCGCGACATCGACCTCAAGATCGCGGGGAGCGCCCTTCGCC 267  
 QY 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68  
 Db 268 TTGTTGATGTTGAGGACCCGCGAGACGCGAGATCGGTGATGTTGCGACGCTAT 327  
 QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaIleGlyArg-----Arg 85  
 Db 328 GATTACGATGGTACCTCTCGGGTGGAGTTTCTCGAAGCGCGCTGGAACAGGCCGA 387  
 QY 86 \*\*\*\*\*SerTyrSerAlaSerArgAlaProSerArg 105  
 Db 388 GCGCGCGGGGTGGAGTGGCGGAGCTCCCGAGTCCGCTATGGCCCCCATCCAGG 447  
 QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125

Db 448 CGGTCTGAAAACAGAGTGGTGTCTCTGGACTGCCCTCAAGTGGAAAGTTGGCAGGATTTA 507  
 QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145  
 Db 508 AAGGATCACATGGTGAAGCAGGTGATGTTATGCTGATGTTTACCAGATGGCACT 567  
 QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165  
 Db 568 GGTGCTCGTGGAGTTTGACGAAAGAGATATACCTATGCAGTTCGAAACTGGATAAC 627  
 QY 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185  
 Db 628 ACTAAGTTTAGATCTCATGAGTAGGTATATACACGATTCTTTCTTTGACCAGAAATTGG 687  
 QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSer 204  
 Db 688 ATACAGTGGTCTTAACACAGTGGAAATTTCAAGGTAAAGATTCCAGGCAAGTTGTCAAGT 744

RESULT 8  
 US-09-770-791-783  
 ; Sequence 783, Application US/09770791  
 ; Patent No. US20020062014A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gorlach, Jorn  
 ; APPLICANT: An, Yong-Qiang  
 ; APPLICANT: Hamilton, Carol M.  
 ; APPLICANT: Price, Jennifer L.  
 ; APPLICANT: Raines, Tracy M.  
 ; APPLICANT: Yu, Yang  
 ; APPLICANT: Rameaka, Joshua G.  
 ; APPLICANT: Page, Amy  
 ; APPLICANT: Matthew, Abraham V.  
 ; APPLICANT: Ledford, Brooke L.  
 ; APPLICANT: Woessner, Jeffrey P.  
 ; APPLICANT: Haas, William David  
 ; APPLICANT: Garcia, Carlos A.  
 ; APPLICANT: Kricker, Maja  
 ; APPLICANT: Slader, Ted  
 ; APPLICANT: Davis, Keith R.  
 ; APPLICANT: Allen, Keith  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Huban, Patrick  
 ; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
 ; FILE REFERENCE: 2029 (PARA-018PRV)  
 ; CURRENT APPLICATION NUMBER: US/09/770,791  
 ; CURRENT FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: 60/178,480  
 ; NUMBER OF SEQ ID NOS: 999  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 783  
 ; LENGTH: 345  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-770-791-783

Alignment Scores:  
 Pred. No.: 3,32e-11 Length: 345  
 Score: 81.00 Matches: 71  
 Percent Similarity: 77.14% Conservative: 10  
 Best Local Similarity: 67.62% Mismatches: 24  
 Query Match: 36.49% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-09-770-791-783 (1-345)  
 QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20  
 Db 4 ATGAGCAGTCTGTCAGTAGAACCGGTAGCTCGGAACCTTCCTGGCATATCCGTGAG 63  
 QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40

```
Db 64 AGAGAGTCGAAGATTGTTTTCAGTAAGTATGGACCTGTTTCAAAATTGATTGAAGGTT 123
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAlaAsp 60
Db 124 CCTCCAAAGGCTCTGTTTATGATTCGTTGATTTGATGATCTCGGATGCTGAAGAT 183
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 184 GCTATTCACTGGTGTGATGCTATGACTTTGATGGGCATCGTTTGAGGTGGAATTGGCG 243
Qy 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 244 CATGTGGGAGCGTTTCAATCAGATGATACTCGGGTAGTTTCAATGGTGGTGGCGGTGT 303
Qy 101 ArgAlaProSerArg 105
Db 304 GGTGGTGGCGCGGT 318
RESULT 9
US-10-060-036-52
; Sequence 52, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-52
Alignment Scores:
Pred. No.: 1,3e-08 Length: 446
Score: 73.50 Matches: 76
Percent Similarity: 65.15% Conservative: 10
Best Local Similarity: 57.58% Mismatches: 41
Query Match: 33.11% Indels: 5
DB: 15 Gaps: 1
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-10-060-036-52 (1-446)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 64 ATCTACCTGGGAGCACTTCGACCGCAGCTGCGGAGAGACTTGGAGACCTGTCTAC 123
Qy 29 LysTyrGlyProIleValAlaPheAspLeuIleProProArgProProGlyTyrAla 48
Db 124 AAGTACGCGCGCATTCGCGAGATCGAGCTCAAGAACCGGCACGCGCTCGCTTCGCC 183
Qy 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
Db 184 TTCGTGGCTTCGAGGACCCCCGAGATGCGAGAGATGCTATTATGGAAGAAATGGTTAT 243
Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyClyArgArg***** 88
Db 244 GATTATGGCCAGTGTGGCTTCGTGTGGAG-----TTCCCGAGACTTAT 288
Qy 89 *****SerTyrSerAlaSerArgAlaProSerArgSerArg 108
Db 289 GGAGGTGGGGTGGGCGGCGGCTGTGGGAGGAATGGCGCTCTCAAGAAGATCTGTAT 348
```

```
Qy 109 TyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHis 128
Db 349 TTCGAGTCTCTGTTTTCAGGACTTCCTCCGTCAGGACGCTGGCAGGACCTGAAGGATCAC 408
Qy 129 MetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 409 ATGCGAAGAGCTGGGGATGCTCTGTTATCTGCTGATGTG 444
RESULT 10
US-09-910-943-651
; Sequence 651, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/1G148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 651
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(759)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-651
Alignment Scores:
Pred. No.: 1.98e-06 Length: 759
Score: 69.00 Matches: 81
Percent Similarity: 42.92% Conservative: 10
Best Local Similarity: 38.21% Mismatches: 117
Query Match: 31.08% Indels: 4
DB: 9 Gaps: 2
US-10-014-927-19MOD_COPY_1_222 (1-222) x US-09-910-943-651 (1-759)
Qy 2 SerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCys 21
Db 109 TCTGGATCCGGCGATGGACGGATATATGTCGGGAATTCGCCGCTGATATTCGGGAGAAG 168
Qy 22 GluValGluAspLeuPheTyrLysTyrGlyProIleValAlaPheAspLeuLys---Ile 40
Db 169 GAAGTACAGAGATCTCTTTGATCGCTATGGTAGGATCCGACCGTAGAGTTGAAGAACCGG 228
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 229 GCGGCGAGTAGTGCCTTCGCTATTCAGCTATCAGGACCCCGCTGATGCAGAGGAC 288
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGlu----- 78
Db 289 GCAGTGTTCGGAAGAAATGGCTATGACTTTGGCTCGTGTGCTACTACGTGTGAGTTTCGG 348
Qy 79 ---IleAlaHisGlyArgArg*****SerTyr 97
Db 349 GTTCTTCGCGGATCTCGCGCGGAGGTGGTGGCGGGGATATGAGGCTCCCGG 408
Qy 98 SerAlaSerArgAlaProSerArgSerAspTyrArgValLeuValThrGlyLeuPro 117
Db 409 GGAAGAAATGGTCCGCCATCTCGCGCTCTGAATACAGAGTCATTGCTCAGGCTCTTCCA 468
Qy 118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
Db 469 CCCTCAGAGAGCTGGCAGGATCTGAGGATCATATGCGGAAGCTGCTGATGCTGTAT 528
Qy 138 SerGluValPheProAspArgLysGlyMetSerClyValValAspTyrSerAsnTyrAsp 157
Db 529 GCTTGATGTACACAAAGATGGAATGGGATGATGCGAATTCATTCCAAAGAGATATGGAA 588
```





; Sequence 38, Application US/10439703  
; Publication No. US20040018527A1  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Jenny  
; APPLICANT: O'Connell, Peter  
; TITLE OF INVENTION: Differential Patterns of Gene Expression that Predict for Docetax  
; TITLE OF INVENTION: Chemosensitivity and Chemoresistance  
; FILE REFERENCE: HO-P248208/10205813  
; CURRENT APPLICATION NUMBER: US/10/439,703  
; PRIOR FILING DATE: 2003-05-16  
; PRIOR APPLICATION NUMBER: US 60/381,141  
; PRIOR FILING DATE: 2002-05-17  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 38  
; LENGTH: 2167  
; TYPE: DNA  
; ORGANISM: Human  
US-10-439-703-38

Alignment Scores:  
Pred. No.: 0.0114 Length: 2167  
Score: 62.00 Matches: 52  
Percent Similarity: 29.25% Conservative: 10  
Best Local Similarity: 24.53% Mismatches: 150  
Query Match: 27.93% Indels: 0  
DB: 12 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-10-439-703-38 (1-2167)

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| QY | 10  | TyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyrLys | 29  |
| DB | 119 | TACATCGCCCGCTGAGTACACAGCCCGGAGCGGATGTGGAGCGCTTCTTTAAAGGC     | 178 |
| QY | 30  | TyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAlaPhe | 49  |
| DB | 179 | TACGGGAAGATCTCGAGGTGGATCTGAAGAACGGATATGGTTTGTGGAGTTTGATGAT   | 238 |
| QY | 50  | ValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAsp    | 69  |
| DB | 239 | CTGGCGTATGCAGATGATGCTGTTTATGAAGTGAATGGCAAGACCTTTGTGTGACGA    | 298 |
| QY | 70  | PheAspGlyCysArgLeuArgValGluIleAlaHisGlyArgArg                | 89  |
| DB | 299 | GTAATTGTGAGCATGCCCGGCCACGGCGAGATGGCAGTTACGGTTCTGGACGCAGT     | 358 |
| QY | 90  | *****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyr              | 109 |
| DB | 359 | CGATATGTTATAGAAGAAAGTGGCGAGATAAATATGCCCTCTCTACTCGCACAGATAC   | 418 |
| QY | 110 | ArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMet | 129 |
| DB | 419 | AGACTTATTGGAGAAATTGTCAGTGGTGCAGCTGGCAAGACCTTAAAGGATTATATG    | 478 |
| QY | 130 | ArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGly | 149 |
| DB | 479 | CGTCAGGACGAGAGTGACTTATGCAGATGCTCACAGGACGCAAAATGAGGGGTG       | 538 |
| QY | 150 | ValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThr | 169 |
| DB | 539 | ATTGAATTGTATCTTATCTGATGATAAAGAGCTTTGGAAAGTTGGATGCAACTGAA     | 598 |
| QY | 170 | GluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSer | 199 |
| DB | 599 | GTCAATGGGAGAAAATCAGATTAGTTGAACAGACCGAGTTCACAGACGCCCGCTCC     | 658 |
| QY | 190 | ValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyPro | 209 |
| DB | 659 | TACTCCAGACCGGAGTCATTCAAGTCTCGCTCTCGAGCAGACATTCCTCGTAGAGC     | 718 |
| QY | 210 | SerCysSerTyrSerSerLysSerArgSerValSer                         | 221 |
| DB | 719 | AGAGCCGAAGTGGCAGCAGCAAGCAAGCATCTTCT                          | 754 |

RESULT 14  
US-10-388-934-197  
; Sequence 197, Application US/10388934  
; Publication No. US20040005547A1  
; GENERAL INFORMATION:  
; APPLICANT: Boess, Franziska  
; APPLICANT: Suter-Dick, Laura  
; APPLICANT: Wolf, Detlef  
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY  
; FILE REFERENCE: 21199  
; CURRENT APPLICATION NUMBER: US/10/388,934  
; CURRENT FILING DATE: 2003-03-14  
; PRIOR APPLICATION NUMBER: 02005336.9  
; PRIOR FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 02015657.6  
; PRIOR FILING DATE: 2002-07-17  
; NUMBER OF SEQ ID NOS: 862  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 197  
; LENGTH: 1579  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)  
US-10-388-934-197

Alignment Scores:  
Pred. No.: 0.0956 Length: 1579  
Score: 57.00 Matches: 47  
Percent Similarity: 26.03% Conservative: 10  
Best Local Similarity: 21.48% Mismatches: 162  
Query Match: 25.68% Indels: 0  
DB: 12 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-10-388-934-197 (1-1579)

|    |     |                                                                |     |
|----|-----|----------------------------------------------------------------|-----|
| QY | 3   | SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu   | 22  |
| DB | 243 | AGTGCTGTCGAGTGTTCATTGGGAGACTAATCCCGCAGCGAGGAGAGATGTGAA         | 302 |
| QY | 23  | ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIlePro      | 42  |
| DB | 303 | AGATTCTTCAGGGTTACGGCGGATCAGATATGATTGTTGAAAAGAGGTTTGGTTT        | 362 |
| QY | 43  | ArgProProGlyTyrAlaPheValGluAspProArgAspAlaAspAlaIle            | 62  |
| DB | 363 | GTGGAAATTGAGGCCCAAGGATCGATGACGCTGTTTATGAATGATGGGAGGAA          | 422 |
| QY | 63  | TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly   | 82  |
| DB | 423 | CTTTCAGTGAAGGGTTACTATTGAACATGCTCGGGCTCGTCTCGAGGTGGAAGGT        | 482 |
| QY | 83  | GlyArgArg                                                      | 102 |
| DB | 483 | AGAGACGACATCTCCGACCGTTTTAGCAGTCGACACCTCGAAATGATAGACGAATGCT     | 542 |
| QY | 103 | ProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyr   | 122 |
| DB | 543 | CCACTCTGAAGACAGAAAATCGACTTATAGTTGAGATTTATCTCAAGAGTCAGCTGG      | 602 |
| QY | 123 | GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro   | 142 |
| DB | 603 | CAGGATCTCAAAAGATTTTCATGAGACACAGCTGGGGAAGTAACCTTTCCGATGCACATCGA | 662 |
| QY | 143 | AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla   | 162 |
| DB | 663 | CCTAAACTAAATGAAGGGGTAGTTGAGTTTCCCTCTTATGTGACTTAAAGATGCTATT     | 722 |
| QY | 163 | IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal   | 182 |
| DB | 723 | GAGAACTTTCTGGAAAGAAATTAATGGGAGAAAATCAAAATTAATTAAGGCGAGCAA      | 782 |
| QY | 183 | ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer   | 202 |

Db 783 AGACACAGGTCAAGAGCAGGTCAAGATCTCGGACCAGGAGTCTCTTAGGTCCCGTAGC 842  
 Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221  
 Db 843 CGATCCCGTTCAGCGAGGACGAAGTCTTACAGCCGATCAAGGAGCAGGAGCGGAGC 899

RESULT 15

US-10-014-927-18  
 ; Sequence 18, Application US/10014927  
 ; Publication No. US20020115180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Barta, Andrea  
 ; APPLICANT: Lopato, Sergey  
 ; APPLICANT: Kalyna, Maria  
 ; APPLICANT: Dornier, Silke  
 ; TITLE OF INVENTION: Splice Factor  
 ; FILE REFERENCE: SONN:013US  
 ; CURRENT APPLICATION NUMBER: US/10/014, 927  
 ; CURRENT FILING DATE: 2001-10-23  
 ; PRIOR APPLICATION NUMBER: PCT/AT00/00100  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: A 727/99  
 ; PRIOR FILING DATE: 1999-04-23  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 18  
 ; LENGTH: 4044  
 ; TYPE: DNA  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of the unknown organism: genome  
 ; OTHER INFORMATION: atSRp30  
 US-10-014-927-18

Alignment Scores:  
 Pred. No.: 6.67 Length: 4044  
 Score: 56.00 Matches: 46  
 Percent Similarity: 26.92% Conservative: 10  
 Best Local Similarity: 22.12% Mismatches: 152  
 Query Match: 25.23% Indels: 0  
 DB: 14 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x US-10-014-927-18 (1-4044)

Qy 11 ValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyrLysTyr 30  
 Db 1659 GTGATTAGGATCACATCGCCAAAGCTGGAGATGCTGCTTCTCTGAAGTTTCCCTGACC 1718  
 Qy 31 GlyProIleValAspIleAspLeuLysIleProProArgProGlyTyrAlaPheVal 50  
 Db 1719 GTAAAGGTGAGTTGACATTCGATGTTTGGATAAGCTTTTTCGATTGATGTTAGTAAT 1778  
 Qy 51 GluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe 70  
 Db 1779 TAGTCTTTGTGAAGGAGATAGTGTGAAGCATCTGAACCTGCTAACTCACATTCAGTAT 1838  
 Qy 71 AspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg\*\*\*\*\* 90  
 Db 1839 TTTCTTTGAGGCATGCTCGGGTTGTGGATTATAGCAACTATGATGATGAAGTACGCA 1898  
 Qy 91 \*\*\*\*\*SerTyrSerAlaSerAlaProSerArgArgSerAspTyrArg 110  
 Db 1899 GTAAAGTTTATATCTTTGGCAACGCAATGTTCTCTGGACTTATGCTTACTGCTTTTGT 1958  
 Qy 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130  
 Db 1959 TTCATAGTATACGACCTGAATTTATCTTCTCGAGGCCAGGTTGGATCTTTGATGTC 2018  
 Qy 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150  
 Db 2019 CCTTAAATTTTGTGTCACAGATAAGGAAACTTGTATGCCACTGAATTCGAAATGCTT 2078  
 Qy 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170

Db 2079 TCCTAGTGTCTATATACGGGTATGTTGTATTGCTTTCTTTGATTGTTTAAAGATAAGT 2138  
 Qy 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190  
 Db 2139 GGATATGGAGTCTATCTCTGAATTTACTGTTTCAGTGCAGGGAATATGATGTCAGGAGTGTG 2198  
 Qy 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210  
 Db 2199 AGTCGAAGCCCGAGATGATTCTAAAGCTATAGAAGCAGGAGTCGGAGCCGTGTGTCGAAGC 2258  
 Qy 211 CysSerTyrSerSerLysSerArg 218  
 Db 2259 TGTAGCTATAGTAGCAAGACAGG 2282

Search completed: February 4, 2004, 21:32:39  
 Job time : 357 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 21:15:14 ; Search time 2041 Seconds  
(without alignments)  
2643.603 Million cell updates/sec

Title: US-10-014-927-19mod\_copy\_1\_222  
Perfect score: 222  
Sequence: 1 MSSRWNTIYVGNLPGDIRK.....RSRSRGSCSYSKSRVSP 222

Scoring table: UNITARY2  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/BAUM927/runat\_04022004\_131334\_2716/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCTL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THRESHOLD=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=BAUM927 -CGN 1 1 2810 @runat\_04022004\_131334\_2716 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estio:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Match | Length | DB | ID       | Description        |
|------------|-------|-------|--------|----|----------|--------------------|
| 1          | 160   | 72.1  | 604    | 9  | AV826310 | AV826310 AV826310  |
| 2          | 156   | 70.3  | 779    | 13 | BQ802976 | BQ802976 WHE2832 C |
| 3          | 154   | 69.4  | 647    | 12 | B1931752 | B1931752 EST551641 |
| 4          | 153   | 68.9  | 877    | 10 | BG444501 | BG444501 GA_Ea002  |
| 5          | 152   | 68.5  | 627    | 10 | BG135885 | BG135885 EST468777 |
| 6          | 151   | 68.0  | 638    | 13 | BQ986483 | BQ986483 HFL1P24r  |
| 7          | 151   | 68.0  | 686    | 10 | BG404250 | BG404250 GA_Ea000  |
| 8          | 150   | 67.6  | 834    | 9  | AW448238 | AW448238 BRY_1489  |
| 9          | 150   | 67.6  | 834    | 13 | BQ605903 | BQ605903 BRY_1489  |
| 10         | 150   | 67.6  | 890    | 10 | BG414461 | BG414461 HVSMEK000 |
| 11         | 148   | 66.7  | 643    | 12 | BW111827 | BW111827 EST559363 |
| 12         | 147   | 66.2  | 852    | 14 | CD439565 | CD439565 EL01N0526 |
| 13         | 146.5 | 66.0  | 1255   | 11 | AY108591 | AY108591 Zea mays  |
| 14         | 146   | 65.8  | 614    | 14 | CA008486 | CA008486 HU11B23r  |
| 15         | 146   | 65.8  | 625    | 12 | BJ269490 | BJ269490 BJ269490  |
| 16         | 146   | 65.8  | 652    | 12 | BQ211859 | BQ211859 BJ211859  |
| 17         | 146   | 65.8  | 687    | 13 | BQ841884 | BQ841884 WHE2986 D |
| 18         | 145   | 65.3  | 631    | 13 | BQ240790 | BQ240790 TaR05012H |
| 19         | 145   | 65.3  | 658    | 12 | BJ314572 | BJ314572 BJ314572  |
| 20         | 145   | 65.3  | 667    | 10 | BG451868 | BG451868 NF101A10D |
| 21         | 144.5 | 65.1  | 840    | 14 | CB650587 | CB650587 OSUNEB15B |
| 22         | 143   | 64.4  | 560    | 10 | BG134853 | BG134853 EST467745 |
| 23         | 143   | 64.4  | 629    | 9  | AI482838 | AI482838 EST242161 |
| 24         | 143   | 64.4  | 666    | 12 | BM878853 | BM878853 P5-G02 Sw |
| 25         | 143   | 64.4  | 736    | 14 | CA802008 | CA802008 sau28e1.0 |
| 26         | 143   | 64.4  | 741    | 14 | CB635325 | CB635325 OS1IEB15J |
| 27         | 143   | 64.4  | 778    | 10 | BZ274242 | BZ274242 GA_EB002  |
| 28         | 142.5 | 64.2  | 810    | 14 | CB662124 | CB662124 OSUNED5N  |
| 29         | 142   | 64.0  | 604    | 12 | BJ209636 | BJ209636 BJ209636  |
| 30         | 142   | 64.0  | 640    | 14 | CA018302 | CA018302 HV08D15r  |
| 31         | 142   | 64.0  | 813    | 14 | CB634299 | CB634299 OS1IEB13L |
| 32         | 141.5 | 63.7  | 677    | 12 | BM063881 | BM063881 KS01060F0 |
| 33         | 141   | 63.5  | 604    | 12 | B1179535 | B1179535 EST520480 |
| 34         | 141   | 63.5  | 657    | 12 | BJ302592 | BJ302592 BJ302592  |
| 35         | 141   | 63.5  | 694    | 10 | BF622642 | BF622642 HVSMEa000 |
| 36         | 141   | 63.5  | 719    | 14 | CA783335 | CA783335 sat22G12  |
| 37         | 140   | 63.1  | 612    | 10 | BE404280 | BE404280 WHE1204 D |
| 38         | 140   | 63.1  | 638    | 13 | BQ238916 | BQ238916 TaE05039F |
| 39         | 140   | 63.1  | 675    | 10 | BG648263 | BG648263 EST509882 |
| 40         | 139   | 62.6  | 677    | 9  | AL508965 | AL508965 AL508965  |
| 41         | 138.5 | 62.4  | 641    | 12 | BM099586 | BM099586 EBea01 SQ |
| 42         | 137.5 | 61.9  | 624    | 10 | BG368538 | BG368538 HVSMEI001 |
| 43         | 137.5 | 61.9  | 667    | 10 | BE420047 | BE420047 WWS02.86R |
| 44         | 137   | 61.7  | 599    | 14 | CA012262 | CA012262 HT04O19r  |
| 45         | 137   | 61.7  | 654    | 12 | BJ446900 | BJ446900 BJ446900  |

#### ALIGNMENTS

RESULT 1  
AV826310  
LOCUS AV826310 RAPL8 Arabidopsis thaliana cdna clone RAPL08-12-105 5',  
DEFINITION AV826310 mRNA linear EST 01-APR-2002  
mRNA sequence.  
ACCESSION AV826310  
VERSION AV826310.1 GI:19868370  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.



US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x BQ802976 (1-779)

```
Qy 1 MetSerArgTrpAenArgThrIleTyrValGlyAsnLeuProGlyAspIleArgIys 20
Db 14 ATGAGTAGGCGCTGGAGCCGACCAATTACGTTGGGAACCTCCAGGGGATATCAGGAG 73
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuIle 40
Db 74 CGGAGGTGGAGGATCTCTTCAAGATATGGACGATAGTTGAAATTGACTTGAAGGTC 133
Qy 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 134 CCCCCAAGGCTCTGGCTTGGTTTGTGTTGAGTTGAGATCCCGCTGATCTGAGAT 193
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 194 GCGATTATGCGCCAGATGCTACAACTTTGATGGAAATCGGCTTAGGTTGAACCTTGA 253
Qy 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 254 CATGGGGGAGGCGCAACTCTTCATCCCTCCGAAACAGCTATGGTGGTGGGAGCGCGT 313
Qy 101 ArgAlaProSerArgSerArgTyrArgValLeuValThrGlyLeuProSerAla 120
Db 314 GGTGGTGTCTCTAGGATACGGAGTATCGTGTCTGTTACTGAGTACTCTTCTTCTGCA 373
Qy 121 SerTrpGlnAspLeuIysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 374 TCATGGCAAGATCTAAGGACCATATAGAAAGCGCTGGTGTGTTGTTCTCTGAGGTG 433
Qy 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 434 TACCGTGAGGGTGGTGTACTATTGGAATTTGTTGATTATACAACTATGATGATGAAG 493
Qy 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 494 TATGCTATAGGAAGCTTGATGATCTGAATTTTAAATAATGCCCTTCTCTCGAGCGCTATA 553
Qy 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 554 AGGTTGAGGAGATGCTGGAAGAACGAGCGGCTCTCTATTTCAGCAGCGGTAGCAGAAC 613
Qy 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 614 CGAAGTGGAGCTACAGCAGAGTCCAGTCCAAAGAAACCAACATCAGCGGTTACGGC 673
Qy 221 Ser 221
Db 674 TCA 676
```

RESULT 3  
BI931752  
LOCUS  
DEFINITION  
EST551641 tomato flower, 8 mm to preanthesis buds Lycopersicon  
esculentum cDNA clone CTC21A15 5' end, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

647 bp mRNA linear EST 18-OCT-2001  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 647)  
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,  
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,  
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
Generation of ESTs from tomato flower tissue, buds 8 mm -  
preanthesis  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University

100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.  
Location/Qualifiers  
source  
1..647  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA96"  
/db\_xref="taxon:4081"  
/clone="CTC21A15"  
/tissue\_type="flower"  
/dev\_stage="buds 8mm to preanthesis"  
/clone\_lib="tomato flower, 8 mm to preanthesis buds"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Cornell University; sequencing: The  
Institute for Genomic Research; flower buds and flowers  
were taken from greenhouse plants (4-8 wks old, TA96).  
They were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."  
BASE COUNT 166 a 122 c 172 g 187 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.72e-27 Length: 647  
Score: 154.00 Matches: 144  
Percent Similarity: 74.40% Conservative: 10  
Best Local Similarity: 69.57% Mismatches: 53  
Query Match: 69.37% Indels: 0  
DB: 12 Gaps: 0  
US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x BI931752 (1-647)  
Qy 4 ArgTrpAenArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23  
Db 27 CGTCTAAGTCGAGTATCTACGTTGGAAATCTTCTCGTGATATTCGGGAGAGAGAAGTA 86  
Qy 24 GluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuIysIleProProArg 43  
Db 87 GAAATTTGTTTAAAGATATGGTCCCATTTGTGAANTGATTGAAAGTCCACCTAGA 146  
Qy 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63  
Db 147 CCACCTGGTTATGCTTCGTAGAGTTTGAAGATCCTCGTGATGCTGATGATGCCATTCGT 206  
Qy 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly 83  
Db 207 GGGGCTGATGGCTATGACTTTGACGGGCATCGCTGGGAGTTGAACCTTGACATGGTGG 266  
Qy 84 ArgArg\*\*\*\*\*SerTyrSerAlaSerArgAlaPro 103  
Db 267 CGAGGATCATCATATGATCGCCACAGTAGTTACAGTAGTCGAGTCGAGTGGACTT 326  
Qy 104 SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123  
Db 327 TCTAGGCGCTCTGACTATCGGTACTGCTCTCTGACTACGCTCTTCTGCTCTATGSCAA 386  
Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143  
Db 387 GACTTGAAGGATCATATGCGACGAGCTGGAGATCTCTGCTTCTCTCAAGTTTCCGAGAT 446  
Qy 144 ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle 163  
Db 447 CQTGACGGTATGAGAGGGGATTGTGGACTATACCAACTATGATGATGATGAGATACGCGATA 506  
Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183  
Db 507 AAGAAACTTGTAGTACTCTCTCTTTCGCAATCAATCTCTCGAGCATATATTAGGTTGAC 566  
Qy 184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg 203  
Db 567 AAGTATGATAGAGCATAGTATTCAGGAGGTCAGTCCAGTCCATCACTTCTAGAGCAGA 626

QY 204 SerArgSerArgGlyProSer 210 877 bp mRNA linear EST 15-MAR-2001  
Db 627 AGGTACTCAAGAGTAGGAGT 647  
RESULT 4  
Bg444501  
LOCUS Bg444501 877 bp mRNA linear EST 15-MAR-2001  
DEFINITION GA\_Ea0024123f Gossypium arboreum 7-10 dpa fiber library Gossypium  
arborescens cDNA clone GA\_Ea0024123f, mRNA sequence.  
ACCESSION Bg444501  
VERSION Bg444501.1 GI:13354153  
KEYWORDS EST.  
SOURCE Gossypium arboreum  
ORGANISM Gossypium arboreum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
REFERENCE 1 (bases 1 to 877)  
AUTHORS Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry  
D., Wood, T.C., Leslie, A. and Wilkins, T.A.  
TITLE An integrated analysis of the genetics, development, and evolution  
of the cotton fiber  
JOURNAL Unpublished  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 556 7288  
Fax: 864 556 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCATATAGGG  
High quality sequence stop: 706.  
Location/Qualifiers  
1..877  
/organism="Gossypium arboreum"  
/mol\_type="mRNA"  
/strain="AKA"  
/cultivar="8400"  
/db\_xref="taxon:29729"  
/clone="GA\_Ea0024123f"  
/tissue\_type="Fibers isolated from bolls harvested 7-10  
dpa"  
/lab\_host="E. coli"  
/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"  
BASE COUNT 218 a 194 c 228 g 236 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,11e-26 Length: 877  
Score: 153.00 Matches: 143  
Percent Similarity: 71.16% Conservative: 10  
Best Local Similarity: 66.51% Mismatches: 62  
Query Match: 68.92% Indels: 0  
DB: 10 Gaps: 0  
US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x BG444501 (1-877)  
QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20  
Db 90 ATGAGTACCGGTTCCAGCAGACACTTTACGTTGGAATCTTCCGGTGATGTCGGCAG 149  
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40  
Db 150 AGGGAAGTGAAGATTCTTTTATAAGTATGTCCTCCATAGTCAAAATGACTTGAAGATT 209  
QY 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60  
Db 210 CCACCAAGCCTCCAGGTATGCAATTTGTCAGTTTGAAGAGTCGAGATGCTGAAGAT 269  
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80

Db 270 GCCATTCGTGTCGTGATGGATATGACTTTGGAGGGCATCGTTTACGGGTTCAACTTGC 329  
QY 81 HisGlyGlyArgArg\*\*\*\*\*SerTyrSerAlaSer 100  
330 CATGTGTCGTGGCGCTCTTCCATAGATCGTCACAGCAGTTATAGTAGTGGCGTGA 389  
QY 101 ArgAlaProSerArgSerAspTyrArgValLeuValThrGlyLeuProSerAla 120  
390 CGTGGACCATCCAGCGCTTGAATATCGGTGCTAGTACTGGATGCCATCTTCTGCT 449  
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140  
450 TCATGGCAGGACCTCAGGATCAGATGGTCCAGCAGGAGATGTTTGTCTCAAGTT 509  
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLys 160  
510 TTCCGTGATGGTAGGGACATACAGGATGTTGGACTACACCACTATGATGATGAAG 569  
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180  
570 TATGCTATTAGAAACATCGATGACTCTGAGTTTCGTAATGCAATTTCTCGGCAATATGTT 629  
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200  
630 CGGGTTAAGGAATATGATTCTAGCGGGGATTCCTTAGAGCCCTAGTGTGGCCGCAICT 689  
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSer 215  
690 CTCCTAAGAGCGGAGCCGAGTAGAGCCGAGCAAGCCGCTGGTGC 734  
RESULT 5  
Bg135885  
LOCUS Bg135885 627 bp mRNA linear EST 31-JAN-2001  
DEFINITION EST468777 tomato crown gall Lycopersicon esculentum cDNA clone  
CT02232 5' sequence, mRNA sequence.  
ACCESSION Bg135885  
VERSION Bg135885.1 GI:12636073  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE 1 (bases 1 to 627)  
AUTHORS van der Hoeven, R., Sun, H., Cho, J., Utterback, T., Hansen, C., Ronning  
C. and Tanksley, S.  
TITLE Generation of ESTs from tomato crown gall tissue  
JOURNAL Unpublished  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
Location/Qualifiers  
1..627  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CT02232"  
/tissue\_type="crown gall"  
/dev\_stage="crown galls from full-grown plants (8 wks old  
/lab\_host="SOLR"  
/clone\_lib="tomato crown gall"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; Four wk old greenhouse plants were stab inoculated  
on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,  
Cornell U.). Galls were allowed to develop for another 4  
wks, when gall tissue was frozen in liquid nitrogen."  
BASE COUNT 159 a 123 c 164 g 181 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.24e-27 Length: 627  
Score: 152.00 Matches: 142  
Percent Similarity: 75.62% Conservative: 10  
Best Local Similarity: 70.65% Mismatches: 49  
Query Match: 68.47% Indels: 0  
DB: 10 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x BG135885 (1-627)

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| Qy | 4   | ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGluVal | 23  |
| Db | 20  | CGTCTAAGTCGAGTACTACGTTGGAAATCTTCTCGTGATATCGGAGAGAAAGTA       | 79  |
| Qy | 24  | GluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIleProProArg       | 43  |
| Db | 80  | GAAGATTGTTTACAAAGTATGTCCTGATGCGAAATGATTGAAGATGCCACCTGA       | 139 |
| Qy | 44  | ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr    | 63  |
| Db | 140 | CCACTGGTATGCGTTCGTAGAGTTGAAGATCTCGTGATGCTGATGATGCTCGT        | 199 |
| Qy | 64  | GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly | 83  |
| Db | 200 | GGGCGTATGCTATGACTTTGACGGCATCGCTTGGAGTTGAAGTTGACATGCTGGG      | 259 |
| Qy | 84  | ArgArg*****SerTyrSerAlaSerArgAlaPro                          | 103 |
| Db | 260 | CGAGGATCATCATATGATCGGCACAGTAGTAGTACAGTACGAGTCGAGTGGACAT      | 319 |
| Qy | 104 | SerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAlaSerTrpGln    | 123 |
| Db | 320 | TCTAGGCGCTGACTATCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG        | 379 |
| Qy | 124 | AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp | 143 |
| Db | 380 | GACTTGAGGATCATATGCGACGAGCTGGAGATGCTGCTTCTCTCAAGTTTCCAGAT     | 439 |
| Qy | 144 | ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIle    | 163 |
| Db | 440 | CGTGACGGTATGAGAGGATGTTGGACTATACCACTATCATGATGATGATGATGATG     | 499 |
| Qy | 164 | ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg | 183 |
| Db | 500 | AAGAACTTGATGACTCTCTGTTTGGCAATCAATCTCTCGAGCATATATTAGGTGGAC    | 559 |
| Qy | 184 | GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArg | 203 |
| Db | 560 | AGTATGATAAGAGGATGATGCTATTCAGGAGTCCAAAGTCCATACATTCATAGACGA    | 619 |
| Qy | 204 | Ser 204                                                      |     |
| Db | 620 | AGT 622                                                      |     |

RESULT 6  
BU986483  
LOCUS  
DEFINITION HF11P24r HF Hordeum vulgare subsp. vulgare cDNA clone HF11P24  
5-PRIME, mRNA sequence.  
ACCESSION  
VERSION BU986483.1 GI:24237429  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
1 (bases 1 to 638)  
Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.  
Barley ESTs from developing seeds  
Unpublished  
JOURNAL  
Contact: Stein Nils  
Molecular Markers Group, Department Genbank

Institute of Plant Genetics and Crop Plant Research (IPK)  
Cortensstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de  
Insert Length: 638 Std Error: 0.00  
Plate: 11 row: p column: 24  
Seq primer: M13rev.

Location/Qualifiers  
1..638  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/db\_xref="GABI:246423"  
/db\_xref="taxon:112509"  
/clone="HF11P24"  
/tissue\_type="caryopsis"  
/dev\_stage="developing caryopsis, 16-25 DAF (days after  
flowering)"  
/lab\_host="XL10-Gold"  
/clone\_lib="HF"

BASE COUNT 153 a 127 c 195 g 163 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.05e-26 Length: 638  
Score: 151.00 Matches: 141  
Percent Similarity: 72.95% Conservative: 10  
Best Local Similarity: 68.12% Mismatches: 56  
Query Match: 68.02% Indels: 0  
DB: 13 Gaps: 0

```

-19MOD_COPY_1_222 (1-222) x BU986483 (1-638)

ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGluVal 23
CGCTGAGCGGACCATTTACGTTGGAACTCCAGGGGATATCAGGAGCGGAGGTG 72
GluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIleProProArg 43
GAGGATCTGTTCTACAGTATGCGGATAGTGGAAATTGACTTGAAGTCCCCCAAG 132
ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63
CCACCTGGCTTTGCTTTTGTGAGTTTGAAGATCCCGTGATCGGAGATGCGATTCAG 192
GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly 83
GGCGAGATGCTACACTTTGATGGGATCGCTCGAGGTGAACTGCACATGGTGGG 252
ArgArg*****SerTyrSerAlaSerArgAlaPro 103
AGGGCTAACTCTTCATCCCTCTTAACAGCCATGCGTGGAGGACGCGGTGGTGTCTC 312
SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGln 123
TCTAGGCATACAGAGTATCGGTGTTCTGCTTACTGACTACCTTTCTTCTGCAATCGGCA 372
AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143
GATCTAAAGGATCATATGAGAAAGCTGGTGAATTTGTTTCTCTGAGTGTATCGTGAG 432
ArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIle 163

```





|                                                                                                                                                             |     |                                                                                                                                                                                                                                                                                        |               |                      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|----------------------|
| MEDLINE                                                                                                                                                     |     | 22478026                                                                                                                                                                                                                                                                               |               |                      |
| PUBMED                                                                                                                                                      |     | 12590341                                                                                                                                                                                                                                                                               |               |                      |
| COMMENT                                                                                                                                                     |     | Contact: Lambrecht M<br>The Arabidopsis Information Resource<br>Carnegie Institution of Washington, Dept. of Plant Biology<br>260 Panama Street, Stanford, CA 94305, USA<br>Tel: 1 650 325 1521 x 251<br>Fax: 1 650 325 3748<br>Email: rhee@acoma.stanford.edu.<br>Location/Qualifiers |               |                      |
| FEATURES                                                                                                                                                    |     | source                                                                                                                                                                                                                                                                                 |               |                      |
|                                                                                                                                                             |     | 1..834                                                                                                                                                                                                                                                                                 |               |                      |
|                                                                                                                                                             |     | /organism="Triticum aestivum"                                                                                                                                                                                                                                                          |               |                      |
|                                                                                                                                                             |     | /mol_type="mRNA"                                                                                                                                                                                                                                                                       |               |                      |
|                                                                                                                                                             |     | /cultivar="Wyuna"                                                                                                                                                                                                                                                                      |               |                      |
|                                                                                                                                                             |     | /db_xref="taxon:4565"                                                                                                                                                                                                                                                                  |               |                      |
|                                                                                                                                                             |     | /tissue_type="endosperm"                                                                                                                                                                                                                                                               |               |                      |
|                                                                                                                                                             |     | /dev_stage="developing endosperm tissue 8, 10 and 12 DPA (days post anthesis)"                                                                                                                                                                                                         |               |                      |
|                                                                                                                                                             |     | /clone_lib="wheat EST endosperm library"                                                                                                                                                                                                                                               |               |                      |
| BASE COUNT                                                                                                                                                  |     | 198 a                                                                                                                                                                                                                                                                                  | 199 c         | 234 g 200 t 3 others |
| ORIGIN                                                                                                                                                      |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Alignment Scores:                                                                                                                                           |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Pred. No.:                                                                                                                                                  |     | 5.84e-26                                                                                                                                                                                                                                                                               | Length:       | 834                  |
| Score:                                                                                                                                                      |     | 150.00                                                                                                                                                                                                                                                                                 | Matches:      | 140                  |
| Percent Similarity:                                                                                                                                         |     | 71.43%                                                                                                                                                                                                                                                                                 | Conservative: | 10                   |
| Best Local Similarity:                                                                                                                                      |     | 66.67%                                                                                                                                                                                                                                                                                 | Mismatches:   | 60                   |
| Query Match:                                                                                                                                                |     | 67.57%                                                                                                                                                                                                                                                                                 | Indels:       | 0                    |
| DB:                                                                                                                                                         |     | 13                                                                                                                                                                                                                                                                                     | Gaps:         | 0                    |
| US-10-014-927-19MOD_COPY_1_222 (1-222) x B0605903 (1-834)                                                                                                   |     |                                                                                                                                                                                                                                                                                        |               |                      |
| QY                                                                                                                                                          | 1   | MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys                                                                                                                                                                                                                           | 20            |                      |
| Db                                                                                                                                                          | 123 | ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG                                                                                                                                                                                                                             | 182           |                      |
| QY                                                                                                                                                          | 21  | CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle                                                                                                                                                                                                                           | 40            |                      |
| Db                                                                                                                                                          | 183 | CGGAGGTGGAGGATCTCTTCAAGATATGAGCGGATAGTTGAAATTCACCTGAAGTC                                                                                                                                                                                                                               | 242           |                      |
| QY                                                                                                                                                          | 41  | ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp                                                                                                                                                                                                                                 | 60            |                      |
| Db                                                                                                                                                          | 243 | CCCCAAGGCGCTCGCTTTGTTGTTGATTTGAAGATCCCGCTGATCTGAAGAT                                                                                                                                                                                                                                   | 302           |                      |
| QY                                                                                                                                                          | 61  | AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla                                                                                                                                                                                                                              | 80            |                      |
| Db                                                                                                                                                          | 303 | GCATTCATGGCGGACGATGCTCAACTGTGATGGGAATCGGCTTANGGTTGAAGTGA                                                                                                                                                                                                                               | 362           |                      |
| QY                                                                                                                                                          | 81  | HisGlyGlyArgArg*****SerTyrSerAlaSer                                                                                                                                                                                                                                                    | 100           |                      |
| Db                                                                                                                                                          | 363 | CATGGCGGAGGCGCAACTCTTCATCCCTCCGAACAGCTATGCTGGGGAGCGCGT                                                                                                                                                                                                                                 | 422           |                      |
| QY                                                                                                                                                          | 101 | ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAla                                                                                                                                                                                                                              | 120           |                      |
| Db                                                                                                                                                          | 423 | GGTGTGTCTCTAGGCATACGGAGTATCGTGTCTGTTACTGGACTACCTTCTCTGCA                                                                                                                                                                                                                               | 482           |                      |
| QY                                                                                                                                                          | 121 | SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal                                                                                                                                                                                                                           | 140           |                      |
| Db                                                                                                                                                          | 483 | TCATGGCAAGATCTGAAGGACCATATGAGAAAGCGCTGATGTTGTTCTCTGAGGTG                                                                                                                                                                                                                               | 542           |                      |
| QY                                                                                                                                                          | 141 | PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys                                                                                                                                                                                                                           | 160           |                      |
| Db                                                                                                                                                          | 543 | TACCGTAGGGTGGTGTACTATTGGAAATTTGATTTATACAACTATGATGATGAAG                                                                                                                                                                                                                                | 602           |                      |
| QY                                                                                                                                                          | 161 | TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle                                                                                                                                                                                                                           | 180           |                      |
| Db                                                                                                                                                          | 603 | TATGCTATAAGGAAGCTTGATGATGATCTGATCTGATTTAAAAATGCCCTTCTCTCAGCGCCTAAT                                                                                                                                                                                                                     | 662           |                      |
| QY                                                                                                                                                          | 181 | ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr                                                                                                                                                                                                                           | 200           |                      |
| Db                                                                                                                                                          | 663 | AGGTAAGGAGGATGCTGGAAAGCAGCGCGCTCTATTACGCGACCGCGTAACAAACC                                                                                                                                                                                                                               | 722           |                      |
| QY                                                                                                                                                          | 201 | ArgSerArgSerArgGlyProSer                                                                                                                                                                                                                                                               | 210           |                      |
| Db                                                                                                                                                          | 723 | CGAAGTGGCAGCTACAGCAGGAGTCCAAGT                                                                                                                                                                                                                                                         | 752           |                      |
| RESULT 9                                                                                                                                                    |     |                                                                                                                                                                                                                                                                                        |               |                      |
| B0605903                                                                                                                                                    |     |                                                                                                                                                                                                                                                                                        |               |                      |
| LOCUS                                                                                                                                                       |     |                                                                                                                                                                                                                                                                                        |               |                      |
| DEFINITION                                                                                                                                                  |     |                                                                                                                                                                                                                                                                                        |               |                      |
| BRY_1489 wheat EST endosperm library Triticum aestivum cDNA 5', mRNA sequence.                                                                              |     |                                                                                                                                                                                                                                                                                        |               |                      |
| ACCESSION                                                                                                                                                   |     |                                                                                                                                                                                                                                                                                        |               |                      |
| B0605903                                                                                                                                                    |     |                                                                                                                                                                                                                                                                                        |               |                      |
| VERSION                                                                                                                                                     |     |                                                                                                                                                                                                                                                                                        |               |                      |
| B0605903.1 GI:21555035                                                                                                                                      |     |                                                                                                                                                                                                                                                                                        |               |                      |
| KEYWORDS                                                                                                                                                    |     |                                                                                                                                                                                                                                                                                        |               |                      |
| EST.                                                                                                                                                        |     |                                                                                                                                                                                                                                                                                        |               |                      |
| SOURCE                                                                                                                                                      |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Triticum aestivum (bread wheat)                                                                                                                             |     |                                                                                                                                                                                                                                                                                        |               |                      |
| ORGANISM                                                                                                                                                    |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Triticum aestivum                                                                                                                                           |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum. |     |                                                                                                                                                                                                                                                                                        |               |                      |
| REFERENCE                                                                                                                                                   |     |                                                                                                                                                                                                                                                                                        |               |                      |
| 1 (bases 1 to 834)                                                                                                                                          |     |                                                                                                                                                                                                                                                                                        |               |                      |
| AUTHORS                                                                                                                                                     |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Clarke, B., Lambrecht, M. and Rhee, S.Y.                                                                                                                    |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Arabidopsis genomic information for interpreting wheat EST sequences                                                                                        |     |                                                                                                                                                                                                                                                                                        |               |                      |
| TITLE                                                                                                                                                       |     |                                                                                                                                                                                                                                                                                        |               |                      |
| Funct. Integr. Genomics 3 (1-2), 33-38 (2003)                                                                                                               |     |                                                                                                                                                                                                                                                                                        |               |                      |
| JOURNAL                                                                                                                                                     |     |                                                                                                                                                                                                                                                                                        |               |                      |

|                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-----------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| MEDLINE                                                   | 22478026                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| PUBMED                                                    | 12590341                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| COMMENT                                                   | Contact: Lambrecht M<br>The Arabidopsis Information Resource<br>Carnegie Institution of Washington, Dept. of Plant Biology<br>260 Panama Street, Stanford, CA 94305, USA<br>Tel: 1 650 325 1521 x 251<br>Fax: 1 650 325 3748<br>Email: rheea@coma.stanford.edu.<br>Location/Qualifiers<br>1. 834<br>/organism="Triticum aestivum"<br>/mol type="mRNA"<br>/cultivar="Wyuana"<br>/db_xref="taxon:4555"<br>/tissue type="endosperm"<br>/dev stage="developing endosperm tissue 8, 10 and 12 DPA (days post anthesis)"<br>/clone lib="wheat EST endosperm library" |
| FEATURES                                                  | BASE COUNT 198 a 199 c 234 g 200 t 3 others<br>ORIGIN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Alignment Scores:                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Pred. No.:                                                | 5.84e-26 Length: 834                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Score:                                                    | 150.00 Matches: 140                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Percent Similarity:                                       | 71.43% Conservative: 10                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| Best Local Similarity:                                    | 66.67% Mismatches: 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Query Match:                                              | 67.57% Indels: 0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| DB:                                                       | 13 Gaps: 0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| US-10-014-927-19MOD_COPY_1_222 (1-222) x BQ605903 (1-834) |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| QY                                                        | 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Db                                                        | 123 ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG 182                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| QY                                                        | 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Db                                                        | 183 CGGAGGTGGAGGATCTCTTCAAGATATGAGCGGATAGTTGAAATTCACCTGAAGTC 242                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| QY                                                        | 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Db                                                        | 243 CCCCCAAGGCGCTCGCTTTGTTGTTGATTTGAAGATCCCGCTGATCTGAAGAT 302                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| QY                                                        | 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Db                                                        | 303 GCATTCATGGCGGACGATGCTCAACTGTGATGGGAATCGGCTTANGGTTGAAGTGA 362                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| QY                                                        | 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Db                                                        | 363 CATGGCGGAGGCGCAACTCTTCATCCCTCCGAACAGCTATGCTGGGGAGCGCGT 422                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| QY                                                        | 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAla 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Db                                                        | 423 GGTGGTGTCTCTAGGCATACGGAGTATCGTGTCTTCTGTTACTGACCTACCTCTCTGCA 482                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| QY                                                        | 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Db                                                        | 483 TCATGGCAAGATCTGAAGGACCATATGAGAAAGCGCTGATGTTCTTCTCTGAGGTG 542                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| QY                                                        | 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Db                                                        | 543 TACCGTAGGGTGGTGTACTATTGGAAATTTGATTTATACAACTATGATGATGAAG 602                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| QY                                                        | 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Db                                                        | 603 TATGCTATAAGGAAGCTTGATGATGATCTGATCTGATTTAAAAATGCCCTTCTCTCAGCGCCTAAT 662                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| QY                                                        | 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Db                                                        | 663 AGGTAAGGAGGATGCTGGAAAGCAGCGCGCTCTCTATTACGCGACCGCGTAACAAACC 722                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| QY                                                        | 201 ArgSerArgSerArgSerArgGlyProSer 210                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

```

||||| |||| | ||||| ||||| |||||
723 CGAAGTGGCAGCTACAGCAGGAGTCCAAGT 752

      890 bp          mRNA       linear    EST 23-OCT-2001
BG414461 HVSMEK0002J05f Hordeum vulgare testa/pericarp EST library
HVCNDA0013 (normal) Hordeum vulgare subsp. vulgare cDNA clone
HVSMEK0002J05f, mRNA sequence.
BG414461
BG414461.1 GI:13320012
EST
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 890)
REFERENCE
Wing R., Close T.J., Kleinhofs A., Wise R., Kannangara G., von
Wetstein D., Akhunov E., Chin A., Choi D.W., Fenton R.D., Kianian
P., Otto C., Simons K., Zhang D., Begum D., Frisch D., Yu Y.,
Henry D., Palmer M., Rambo T., Simmons J., Oates R. and Main D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex testa/pericarp cDNA library
Unpublished
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 511
Seg primer: AATTAACTCTACTAAAGG
High quality sequence stop: 683.
Location/Qualifiers
1. 890
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Vorex"
/db_xref="taxon:112509"
/clone="HVSMEK0002J05f"
/tissue_type="testa/pericarp"
/lab_host="rjci21"
/clone_lib="Hordeum vulgare testa/pericarp EST library
HVCNDA0013 (normal)"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Plants were raised from seeds in a Controlled Environments
growth chamber maintained in continuous light at 18oC, and the
testa and pericarp were dissected from developing kernels
at Washington State University. Pullman, WA (Kannangara,
von Wetstein). Total RNA was prepared, poly(A) RNA was
purified, one cDNA library was made, and 1 million pfu
were in vivo excised to give pluescript SK(-) cDNA
phagemids in the TJ Clouse lab at the University of
California, Riverside (Akunov, Chin, Choi, Close, Fenton,
Kianian, Otto, Simons, Zhang). Phagemids were plated and
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
189 c 231 g 244 t 1 others
225 a

```

JOURNAL  
COMMENT

Unpublished  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@igr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T3.

FEATURES  
source

1. .643  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="CPRO12D7"  
/tissue\_type="roots"  
/dev\_stage="in vitro grown stem cuttings"  
/lab\_host="SOLR"  
/clone\_lib="potato roots"  
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Cornell University, Tanksley lab,  
sequencing; The Institute for Genomic Research. Roots were  
isolated from in vitro grown stem cuttings on CM medium.  
Roots were isolated two weeks after placing the stem  
cuttings from in vitro grown plants on medium."  
BASE COUNT 159 a 124 c 167 g 193 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,97e-26 Length: 643  
Score: 148.00 Matches: 138  
Percent Similarity: 74.37% Conservative: 10  
Best Local Similarity: 69.35% Mismatches: 51  
Query Match: 66.67% Indels: 0  
DB: 12 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x BM111827 (1-643)

QY 4 ArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23  
DB 47 CGTCTAAGTCGGACTATCTACGTCGGAATCTCTCGGTGATATTCGGGAGAGAAATG 106  
QY 24 GluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProArg 43  
DB 107 GAAGATTGTTTACAAGTATGTCCTCCATTTGGAAATGATTGAAATTCACCTAGA 166  
QY 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaIleTyr 63  
DB 167 CCACCTGGTTATCGGTTCTAGAGTTTGAAGATCCTCGTGATCTGATGATGCCATCCGT 226  
QY 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 83  
DB 227 GGCGCGTGGTGGCTATGACTTTGATGGCGTCGCTTGGAGTTGAACTTGCACATGGTGG 286  
QY 84 ArgArg\*\*\*\*\*SerTyr:SerAlaSerArgAlaPro 103  
DB 287 CGAGGATCATCATATATGATCCACAGTATGACGTAGTGGAGTGCATGGTGGATT 346  
QY 104 SerArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyrGln 123  
DB 347 TCTAGCGGCTCTGACTATCGGTACTGGTCTCTGGACTACCATCTCTGCTTCATGGCAA 406  
QY 124 AspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAsp 143  
DB 407 GACTTGAAGGATCACATGCGAGCGAGTGGAGATGTCTGCTTCTCAAGTTTTCGGAGAT 466  
QY 144 ArgLysGlyMet:SerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIle 163  
DB 467 CGTGACGGATGAGAGGGATTGGGACTATACCACTATGATGATGATGATGATGATGATGAT 526  
QY 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183

Db  
QY

527 AAGAACTTGATGACTCTCTGTTTCGCAATCATCTCTCGAGCATATATTAGGTGGGA 586  
184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202  
DB 587 TAGTATGATAAGAGGATGATCTATTCAGGAGTCCAGTCCATATATTATTCAGAAGC 643

RESULT 12  
CD439565

LOCUS CD439565 852 bp mRNA linear EST 03-JUN-2003  
DEFINITION ELO1N0526D06.b Endosperm\_5 Zea mays cDNA, mRNA sequence.  
ACCESSION CD439565  
VERSION CD439565.1 GI:31355208  
KEYWORDS EST.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 852)  
AUTHORS La,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and  
Messing,J.  
TITLE Sequencing of the maize endosperm ESTs  
JOURNAL Unpublished  
COMMENT Contact: Lai, Jinheng  
Dr. Joachim Messing's lab  
Waksman Institute, Rutgers University  
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
Tel: 732-445-3801  
Fax: 732-445-5735  
Email: jlai@waksman.rutgers.edu  
Seq primer: T3.

FEATURES  
source

Location/Qualifiers  
1. .852  
/organism="Zea mays"  
/mol\_type="mRNA"  
/cultivar="W22"  
/db\_xref="taxon:4577"  
/tissue\_type="Endosperm of 7-23DAP"  
/clone\_lib="Endosperm\_5"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"  
BASE COUNT 217 a 200 c 225 g 210 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.11e-25 Length: 852  
Score: 147.00 Matches: 137  
Percent Similarity: 68.69% Conservative: 10  
Best Local Similarity: 64.02% Mismatches: 67  
Query Match: 66.22% Indels: 0  
DB: 14 Gaps: 0

US-10-014-927-19MOD\_COPY\_1\_222 (1-222) x CD439565 (1-852)

QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20  
DB 75 ATGACCGGGGACGCGCTCTACGATCTAGTGGGACCTCCCGCGACATCCGGAG 134  
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40  
DB 135 AGGGAAGTGGATGATCTCTTCTACAGTATGGAGTATAGTGGAAATTCACCTTGAATAAT 194  
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60  
DB 195 CCACCAAGGCTCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 254  
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80  
DB 255 GCATATATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314  
QY 81 HisGlyGlyArgArg\*\*\*\*\*SerTyr:SerAlaSer 100  
DB 315 CATGTGGAGGACACATCTCTTTTGTGATCGATCTAGCAGCTATAGCAGTCTGCTGGACAA 374

|                   |            |                                                              |                                     |
|-------------------|------------|--------------------------------------------------------------|-------------------------------------|
| QY                | 101        | ArgAlaProSerArgSerArgSerArgValLeuValThrGlyLeuProSerAla       | 120                                 |
| Db                | 375        | CGCGTGCGCTCAAAACGTTTCATACCGTGTATGTTACTGGATTACCTTTCAGCA       | 434                                 |
| QY                | 121        | SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal | 140                                 |
| Db                | 435        | TCATGGCAAGATCTCAAGGACCATATCGCGGAGCTGGTGATGTCGTCTTTCACTGATGTG | 494                                 |
| QY                | 141        | PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys | 160                                 |
| Db                | 495        | TATCGTAGGCTGGAGCAACTATTGGATAGCTGATTATATACTAACTATGAAGATATGAAA | 554                                 |
| QY                | 161        | TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle | 180                                 |
| Db                | 555        | CACGCGATAGAGAGCTAGATGTTCTGAGTTCGTAATGCTTTTTCAGGACATATGTC     | 614                                 |
| QY                | 181        | ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr | 200                                 |
| Db                | 615        | CGGGTGGAGGATGATGCTAGCGGAGCGGCTCTCGCTCCAGAGCGCAACCGCTCT       | 674                                 |
| QY                | 201        | ArgSerArgSerArgSerArgGlyProSerCysSerTyrSer                   | 214                                 |
| Db                | 675        | AAGTCAAGAGAGAGGACAGGACAGCCACTGCTGCTCAAGAGC                   | 716                                 |
| RESULT 13         |            |                                                              |                                     |
| AY108591          |            |                                                              |                                     |
| LOCUS             | AY108591   | Zea mays PCO125156 mRNA sequence.                            | 1255 bp mRNA linear HTC 17-OCT-2002 |
| DEFINITION        | AY108591   |                                                              |                                     |
| ACCESSION         | AY108591   |                                                              |                                     |
| VERSION           | AY108591.1 | GI:21211708                                                  |                                     |
| KEYWORDS          | HTC.       |                                                              |                                     |
| SOURCE            |            |                                                              |                                     |
| ORGANISM          |            |                                                              |                                     |
| REFERENCE         |            |                                                              |                                     |
| AUTHORS           |            |                                                              |                                     |
| TITLE             |            |                                                              |                                     |
| JOURNAL           |            |                                                              |                                     |
| REFERENCE         |            |                                                              |                                     |
| AUTHORS           |            |                                                              |                                     |
| TITLE             |            |                                                              |                                     |
| JOURNAL           |            |                                                              |                                     |
| COMMENT           |            |                                                              |                                     |
| FEATURES          |            |                                                              |                                     |
| SOURCE            |            |                                                              |                                     |
| BASE COUNT        |            |                                                              |                                     |
| ORIGIN            |            |                                                              |                                     |
| Alignment Scores: |            |                                                              |                                     |

```

TITLE      Barley ESTs from germinating seeds
JOURNAL    Unpublished
COMMENT     Contact: Stein Nils
           Molecular Markers Group, Department Genbank
           Institute of Plant Genetics and Crop Plant Research (IPK)
           Corrensstr. 3, 06466, Gatersleben, Germany
           Tel: 039482-5522
           Fax: 039482-5595
           Email: stein@ipk-gatersleben.de
           Insert Length: 614 Std Error: 0.00
           Plate: 11 row: B column: 23
           Seq primer: M13rev.
           Location/Qualifiers
             1..614
               /organism="Hordeum vulgare subsp. vulgare"
               /mol_type="mRNA"
               /cultivar="barke"
               /db_xref="GABI:266136"
               /db_xref="taxon:112509"
               /clone="HUI1B23"
               /tissue_type="germinating seeds"
               /dev_stage="germinating seeds, 16-48 h"
               /lab_host="XL10-Gold"
               /clone_lib="HU"
               /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
               cDNA); Site 2: XhoI (3'-end of cDNA); germinating seeds,
               16-48h. Due to a cloning artefact caused by the kit, in
               most cases the EcoRI site is NOT present, as well as the
               EcoRI adapter used for cloning. To excise the insert,
               restriction sites upstream EcoRI should be used (e.g.
               BamHI, SalI, PstI). NOTE: Also due to the cloning system
               used Blue/white selection for recombinants is not 100%
               reliable. Average insert size is 1.2 kb"
             146 a 118 c 199 g 161 t

BASE COUNT 146 a 118 c 199 g 161 t
ORIGIN
Alignment Scores:
  Pred. No.:      2e-25      Length:      614
  Score:          146.00     Matches:     136
  Percent Similarity: 77.25%   Conservative: 10
  Best Local Similarity: 71.96% Mismatches:    43
  Query Match:     65.77%     Indels:      0
  DB:              14       Gaps:          0

US-10-014-927-19MOD_COPY_1_222 (1-222) x CA008486 (1-614)

QY      1 MetSerSerArgTrpAsnArgThrIleTyValGlyAsnLeuProGlyAspIleArgLys 20
Db      48 ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG 107

QY      21 CysGluValGluAspLeuPheTyLysTyGlyProIleValAspIleAspLeuIle 40
Db      108 CGGGAGGTGGAGGATCTGTCTACAGATATGACCGGATAGTGGAAATGACTTGAAGTTC 167

QY      41 ProProArgProGlyTyxAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db      168 CCCCCAGCCACCTGGCTTCTGTTGTTGATTTGAGATCCCGGTGATGCCGAGAT 227

QY      61 AlaIleTyGlyArgAspGlyTyxAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db      228 GCGATTACGGCGGAGATGCTTACAACTTTGATGGGAATCGCTGAGGTTGAACTTCCA 287

QY      81 HisGlyGlyArgArg*****SerTySerAlaSer 100
Db      288 CATGTGGGAGGCGTAACTCTTCATCCCTTCTAAACAGCCATGTGGTGGAGGAGCGCGT 347

QY      101 ArgAlaProSerArgSerAspTyxArgValLeuValThrGlyLeuProProSerAla 120
Db      348 GGTGGGTCTCTAGGCATACAGAGTATCGTGTCTCGGTACTGGACTTCCTTCTTCTCA 407

QY      121 SerTrpGlnAspLeuIleAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db      408 TCATGCCAGATCTAAAGGATCATATGAGAAGGCTGGTGTATGTTTCTCTGAGGTG 467

TITLE      Barley ESTs from germinating seeds
JOURNAL    Unpublished
COMMENT     Contact: Stein Nils
           Molecular Markers Group, Department Genbank
           Institute of Plant Genetics and Crop Plant Research (IPK)
           Corrensstr. 3, 06466, Gatersleben, Germany
           Tel: 039482-5522
           Fax: 039482-5595
           Email: stein@ipk-gatersleben.de
           Insert Length: 614 Std Error: 0.00
           Plate: 11 row: B column: 23
           Seq primer: M13rev.
           Location/Qualifiers
             1..614
               /organism="Hordeum vulgare subsp. vulgare"
               /mol_type="mRNA"
               /cultivar="barke"
               /db_xref="GABI:266136"
               /db_xref="taxon:112509"
               /clone="HUI1B23"
               /tissue_type="germinating seeds"
               /dev_stage="germinating seeds, 16-48 h"
               /lab_host="XL10-Gold"
               /clone_lib="HU"
               /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
               cDNA); Site 2: XhoI (3'-end of cDNA); germinating seeds,
               16-48h. Due to a cloning artefact caused by the kit, in
               most cases the EcoRI site is NOT present, as well as the
               EcoRI adapter used for cloning. To excise the insert,
               restriction sites upstream EcoRI should be used (e.g.
               BamHI, SalI, PstI). NOTE: Also due to the cloning system
               used Blue/white selection for recombinants is not 100%
               reliable. Average insert size is 1.2 kb"
             146 a 118 c 199 g 161 t

BASE COUNT 146 a 118 c 199 g 161 t
ORIGIN
Alignment Scores:
  Pred. No.:      2e-25      Length:      614
  Score:          146.00     Matches:     136
  Percent Similarity: 77.25%   Conservative: 10
  Best Local Similarity: 71.96% Mismatches:    43
  Query Match:     65.77%     Indels:      0
  DB:              14       Gaps:          0

US-10-014-927-19MOD_COPY_1_222 (1-222) x CA008486 (1-614)

QY      1 MetSerSerArgTrpAsnArgThrIleTyValGlyAsnLeuProGlyAspIleArgLys 20
Db      48 ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG 107

QY      21 CysGluValGluAspLeuPheTyLysTyGlyProIleValAspIleAspLeuIle 40
Db      108 CGGGAGGTGGAGGATCTGTCTACAGATATGACCGGATAGTGGAAATGACTTGAAGTTC 167

QY      41 ProProArgProGlyTyxAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db      168 CCCCCAGCCACCTGGCTTCTGTTGTTGATTTGAGATCCCGGTGATGCCGAGAT 227

QY      61 AlaIleTyGlyArgAspGlyTyxAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db      228 GCGATTACGGCGGAGATGCTTACAACTTTGATGGGAATCGCTGAGGTTGAACTTCCA 287

QY      81 HisGlyGlyArgArg*****SerTySerAlaSer 100
Db      288 CATGTGGGAGGCGTAACTCTTCATCCCTTCTAAACAGCCATGTGGTGGAGGAGCGCGT 347

QY      101 ArgAlaProSerArgSerAspTyxArgValLeuValThrGlyLeuProProSerAla 120
Db      348 GGTGGGTCTCTAGGCATACAGAGTATCGTGTCTCGGTACTGGACTTCCTTCTTCTCA 407

QY      121 SerTrpGlnAspLeuIleAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db      408 TCATGCCAGATCTAAAGGATCATATGAGAAGGCTGGTGTATGTTTCTCTGAGGTG 467

QY      141 PheProAspArgLysGlyMetSerGlyValValAspTyxSerAsnTyxAspMetLys 160
Db      468 TATCGTGAGGGTGGTGGTACTATTGGAATTTGATATACAACTATGATATATGAG 527

QY      161 TyxAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyxIle 180
Db      528 TATGCTATAGGAAGCTTGATGATATGATATTAATAATGCTTCTCTCGAGCGCCTATA 587

QY      181 ArgValArgGluTyxGluSerArgSer 189
Db      588 AGGGTGAAGGAGTATGCTGGCAAAAGC 614

RESULT 15
BU269490 625 bp mRNA linear EST 09-APR-2002
LOCUS     aestivum cDNA clone whoh22d09 5', mRNA sequence.
DEFINITION
ACCESSION BU269490
VERSION    BU269490.1 GI:20095590
KEYWORDS   EST.
SOURCE     Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum
REFERENCE  1 (bases 1 to 625)
AUTHORS   Ogihara,Y. and Murai,K.
TITLE     Expressed genes in Triticum aestivum
COMMENT   Contact: Tadasu Shin-i
           Center For Genetic Resource Information
           National Institute of Genetics
           1111 Yata, Mishima, Shizuoka 411-8540, Japan
           Tel: 81-559-81-6856
           Fax: 81-559-81-6855
           Email: tshini@genes.nig.ac.jp.
           Location/Qualifiers
             1..625
               /organism="Triticum aestivum"
               /mol_type="mRNA"
               /cultivar="Chinese Spring"
               /db_xref="taxon:4565"
               /clone="whoh22d09"
               /tissue_type="pistil at heading date"
               /dev_stage="Peaks", scale 10.5"
               /clone_lib="Y. Ogihara unpublished cDNA library, Wh_oh"
             148 a 126 c 190 g 161 t

BASE COUNT 148 a 126 c 190 g 161 t
ORIGIN
Alignment Scores:
  Pred. No.:      2.15e-25   Length:      625
  Score:          146.00     Matches:     136
  Percent Similarity: 77.25%   Conservative: 10
  Best Local Similarity: 71.96% Mismatches:    43
  Query Match:     65.77%     Indels:      0
  DB:              12       Gaps:          0

US-10-014-927-19MOD_COPY_1_222 (1-222) x BU269490 (1-625)

QY      1 MetSerSerArgTrpAsnArgThrIleTyValGlyAsnLeuProGlyAspIleArgLys 20
Db      54 ATGAGTAGGCGCTGGAGCGGACCATTTACGTTGGGAACCTCCAGGGGATATCAGGGAG 113

QY      21 CysGluValGluAspLeuPheTyLysTyxGlyProIleValAspIleAspLeuIle 40
Db      114 CGGGAGGTGGAGGATCTCTTACAAAGTATGGACGATAGTTGAAATTCACITGAAGTTC 173

QY      41 ProProArgProGlyTyxAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db      174 CCCCCAGGCGCTCTGGCTTGTCTTTGTGATTTGAAGATCCCGGTGATGCTGAAGAT 233

QY      61 AlaIleTyGlyArgAspGlyTyxAspPheAspGlyCysArgLeuArgValGluIleAla 80

```

```

Db      234 GCGATTTCATGCCCGAGATGGTACAACTTTGATGGGAATCGGCTTAGGTTGAACCTTGCA 293
QY      81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db      294 CATGCCGGAGGGCCAACTCTTCATCCCTCCGAACAGCTATGGTGGGGGACGCCGT 353
QY      101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
Db      354 GGTGGTGTCTCTAGGCATACCGAGTATCGTGTCTGCTTACTGGACTACCTTCTCTGCA 413
QY      121 SerTyrGlnAspLeuLysAspHisMetArgLysValGlyAspValCysPheSerGluVal 140
Db      414 TCATGCCCAAGATCTTAAGGACCATATAGAAAGGCTGGTGTATGTTGTTCTCTGAGGTG 473
QY      141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLys 160
Db      474 TACCGTAGGGTGGTGGTACTATTGGAATTGTTGATTATACAAACTATGATGATGAAG 533
QY      161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db      534 TATGCTATAAGGAAGCTTGTATGATGATGATACTGAATTTAAATGCGCTTCTCTCGAGCGCCTATA 593
QY      181 ArgValArgGluTyrGluSerArgSer 189
Db      594 AGGGTGAAGGAGTATGCTGGAAAAAGC 620
```

Search completed: February 4, 2004, 22:08:06  
Job time : 2049 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 15:22:45 ; Search time 14428 Seconds  
(without alignments)  
1466.491 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaagatgcttactctt.....atgattatcttaactaac 4044

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_on.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_to.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rnd.\*

36: em\_htg\_nam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Match | Length | DB | ID        | Description         |
|------------|--------|-------|--------|----|-----------|---------------------|
| 1          | 4044   | 100.0 | 4044   | 6  | AX040661  | AX040661 Sequence   |
| 2          | 4021   | 99.4  | 5164   | 8  | ATH131214 | ATH131214 Arabidops |
| 3          | 3493.2 | 86.4  | 59261  | 8  | T12M4     | AC003114 Arabidops  |
| 4          | 780    | 19.3  | 780    | 6  | AX509198  | AX509198 Sequence   |
| 5          | 571    | 14.1  | 110514 | 8  | F7G19     | AC000106 Sequence   |
| 6          | 311.4  | 7.7   | 933    | 8  | AY050912  | AY050912 Arabidops  |
| 7          | 139    | 3.4   | 109465 | 8  | AP006409  | AP006409 Lotus jap  |
| 8          | 131.4  | 3.2   | 133656 | 8  | AC144541  | AC144541 Medicago   |
| 9          | 117    | 2.9   | 838    | 8  | AY150486  | AY150486 Arabidops  |
| 10         | 113.8  | 2.8   | 939    | 8  | AY128356  | AY128356 Arabidops  |
| 11         | 108.8  | 2.7   | 762    | 6  | AX506504  | AX506504 Sequence   |
| 12         | 107.6  | 2.7   | 120374 | 8  | AC126791  | AC126791 Medicago   |
| 13         | 107.6  | 2.7   | 129199 | 2  | AC134522  | AC134522 Medicago   |
| 14         | 105.4  | 2.6   | 80367  | 8  | AC012329  | AC012329 Arabidops  |
| 15         | 105.4  | 2.6   | 104204 | 8  | ATT9C5    | AL132964 Arabidops  |
| 16         | 98.2   | 2.4   | 107367 | 2  | AC136288  | AC136288 Medicago   |
| 17         | 86.6   | 2.1   | 168267 | 8  | AP003535  | AP003535 Oryza sat  |
| 18         | 78.6   | 1.9   | 5285   | 6  | AR204400  | AR204400 Sequence   |
| 19         | 78.6   | 1.9   | 94536  | 8  | T14P8     | AF069298 Arabidops  |
| 20         | 78.6   | 1.9   | 194892 | 8  | ATCHRIV6  | AL161494 Arabidops  |
| 21         | 78     | 1.9   | 157248 | 8  | AP004300  | AP004300 Oryza sat  |
| 22         | 77.8   | 1.9   | 93731  | 2  | AC138073  | AC138073 Homo sapi  |
| 23         | 75.8   | 1.9   | 7218   | 6  | I66494    | I66494 Sequence 14  |
| 24         | 75     | 1.9   | 66993  | 2  | AC138074  | AC138074 Homo sapi  |
| 25         | 70.4   | 1.7   | 761    | 11 | CNS06KKG  | AL402998 T7 end of  |
| 26         | 67.4   | 1.7   | 6644   | 6  | E23356    | E23356 Virus vecto  |
| 27         | 67.4   | 1.7   | 7372   | 6  | E23357    | E23357 Virus vecto  |
| 28         | 67.4   | 1.7   | 7797   | 6  | E23355    | E23355 Virus vecto  |
| 29         | 67.4   | 1.7   | 7996   | 6  | E23359    | E23359 Virus vecto  |
| 30         | 67.4   | 1.7   | 80216  | 2  | AC139363  | AC139363 Homo sapi  |
| 31         | 67     | 1.7   | 289973 | 2  | AC135678  | AC135678 Rattus no  |
| 32         | 65.4   | 1.6   | 4146   | 8  | AF001035  | AF001035 Arabidops  |
| 33         | 65.4   | 1.6   | 6161   | 6  | AX345313  | AX345313 Sequence   |
| 34         | 65.4   | 1.6   | 100239 | 8  | AC009525  | AC009525 Arabidops  |
| 35         | 65.2   | 1.6   | 286208 | 2  | AC117140  | AC117140 Rattus no  |
| 36         | 64.8   | 1.6   | 14006  | 6  | AX346860  | AX346860 Sequence   |
| 37         | 64.8   | 1.6   | 93975  | 2  | AC139011  | AC139011 Homo sapi  |
| 38         | 64.4   | 1.6   | 6171   | 6  | AX345690  | AX345690 Sequence   |
| 39         | 64.2   | 1.6   | 127709 | 2  | CNS05TCL  | AL352983 Homo sapi  |
| 40         | 64.2   | 1.6   | 156550 | 2  | AC015830  | AC015830 Homo sapi  |
| 41         | 64     | 1.6   | 151368 | 2  | AC142664  | AC142664 Macaca mu  |
| 42         | 63.8   | 1.6   | 143585 | 2  | AC013349  | AC013349 Homo sapi  |
| 43         | 63.2   | 1.6   | 6668   | 6  | AX346599  | AX346599 Sequence   |
| 44         | 63     | 1.6   | 83440  | 2  | AC024285  | AC024285 Homo sapi  |
| 45         | 62.8   | 1.6   | 65034  | 2  | AC016067  | AC016067 Homo sapi  |

ALIGNMENTS

RESULT 1  
AX040661  
LOCUS AX040661 4044 bp DNA linear PAT 23-NOV-2000  
DEFINITION Sequence 18 from Patent WO0065059.  
ACCESSION AX040661  
VERSION AX040661.1 GI:11340389  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1  
AUTHORS Barta, A., Lopato, S., Kalya, M. and Dörner, S.  
TITLE Splicing factor  
JOURNAL Patent: WO 0065059-A 18 02-NOV-2000;  
(sterreichisches Forschungszentrum Seibersdorf Ges.m.b.H.; (AT)



| FEATURES   | Location/Qualifiers                                                     | source |
|------------|-------------------------------------------------------------------------|--------|
|            | 1. .4044                                                                |        |
|            | /organism="unidentified"                                                |        |
|            | /mol_type="genomic DNA"                                                 |        |
|            | /db_xref="taxon:3264"                                                   |        |
|            | /note="Beschreibung des unbekannten ORGANISM: Genom                     |        |
| BASE COUNT | 1061 a 727 c 829 g 1427 t                                               |        |
| ORIGIN     |                                                                         |        |
|            | Query Match 100.0%; Score 4044; DB 6; Length 4044;                      |        |
|            | Best Local Similarity 100.0%; Pred. No. 0;                              |        |
|            | Matches 4044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;           |        |
| Qy         | 1 AGACAAAGATGCTTACTCTTAAACATGTCGAGCTTTATTGAAATGATCACACAGCTTC 60         |        |
| Db         | 1 AGACAAAGATGCTTACTTCTTAAACATGTCGAGCTTTATTGAAATGATCACACAGCTTC 60        |        |
| Qy         | 61 TAACTATGATGTCCTTCCCTGCAAGCGAACAGTGGAAAGATGATTGATAACGGAAATAT 120      |        |
| Db         | 61 TAACTATGATGTCCTTCCCTGCAAGCGAACAGTGGAAAGATGATTGATAACGGAAATAT 120      |        |
| Qy         | 121 CGGAACATCACTCAACCAACCAAAATTTGGACATCATATCGCAACAAATCAATAGGAA 180      |        |
| Db         | 121 CGGAACATCACTCAACCAACCAAAATTTGGACATCATATCGCAACAAATCAATAGGAA 180      |        |
| Qy         | 181 AAATACTGAAATTCACAAACAGAAAACCAACCGAACAGACGAGGAACTCACCGACTGA 240      |        |
| Db         | 181 AAATACTGAAATTCACAAACAGAAAACCAACCGAACAGACGAGGAACTCACCGACTGA 240      |        |
| Qy         | 241 GAGAGACCGTGGACCGGTGFCACGGCGGAAAATGGTCTTGGACGGAGTTACTTAATCGGCGA 300  |        |
| Db         | 241 GAGAGACCGTGGACCGGTGFCACGGCGGAAAATGGTCTTGGACGGAGTTACTTAATCGGCGA 300  |        |
| Qy         | 301 ATTGAGATTTGAGAGTGTGTAGAGAACCGAGAGAAATGTTCTCTCAAAAAAATCCC 360        |        |
| Db         | 301 ATTGAGATTTGAGAGTGTGTAGAGAACCGAGAGAAATGTTCTCTCAAAAAAATCCC 360        |        |
| Qy         | 361 CAAGTGTTCGGATCTAGTGTCTCTTTTGTCCAAAACGACAGTGTTTAGGAAACCTPAGG 420     |        |
| Db         | 361 CAAGTGTTCGGATCTAGTGTCTCTTTTGTCCAAAACGACAGTGTTTAGGAAACCTPAGG 420     |        |
| Qy         | 421 AGAATGAATGACCGGATGTCGGAATCCGATTCGAAATTTGGTTCGAAATTTGAATACTAA 480    |        |
| Db         | 421 AGAATGAATGACCGGATGTCGGAATCCGATTCGAAATTTGGTTCGAAATTTGAATACTAA 480    |        |
| Qy         | 481 CATACAAATATCCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACATAT 540      |        |
| Db         | 481 CATACAAATATCCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACATAT 540      |        |
| Qy         | 541 ATTACAGGCCCATTAACACATATGGCGGATCTTTGATCAACTGGCTATTTCATCGTTGA 600     |        |
| Db         | 541 ATTACAGGCCCATTAACACATATGGCGGATCTTTGATCAACTGGCTATTTCATCGTTGA 600     |        |
| Qy         | 601 TACATCGCGCGCACAGGATTAATAATCCAGTTTCGGTTTTTATAAAGGATACATAGTTTCCA 660  |        |
| Db         | 601 TACATCGCGCGCACAGGATTAATAATCCAGTTTCGGTTTTTATAAAGGATACATAGTTTCCA 660  |        |
| Qy         | 661 AACGAACGGTGGTGTCTCTTCCAGAACAAATCTAACAGCTTCTTCGAAACATCTTCTTCT 720    |        |
| Db         | 661 AACGAACGGTGGTGTCTCTTCCAGAACAAATCTAACAGCTTCTTCTTCGAAACATCTTCTTCT 720 |        |
| Qy         | 721 TCTTCTTCTCGAAATATTTTCCAGTAATCAATTTCTTCTCTCTAGATTTTACAGG 780         |        |
| Db         | 721 TCTTCTTCTCGAAATATTTTCCAGTAATCAATTTCTTCTCTCTAGATTTTACAGG 780         |        |
| Qy         | 781 AACTAATTTTCTGCTCTGAGGTATCAGATGAGTAGCCGATGGAATCGTACGATCTACGTT 840    |        |
| Db         | 781 AACTAATTTTCTGCTCTGAGGTATCAGATGAGTAGCCGATGGAATCGTACGATCTACGTT 840    |        |
| Qy         | 841 GGAATTTCCCTGGAGATATTCGCAAGTGTGAGGTGGAAGATCTCTTCTCAAGGTTGA 900       |        |
| Db         | 841 GGAATTTCCCTGGAGATATTCGCAAGTGTGAGGTGGAAGATCTCTTCTCAAGGTTGA 900       |        |

|    |      |                                                                 |      |
|----|------|-----------------------------------------------------------------|------|
| QY | 1981 | TTATCTTCTCGAGGCGCAGTGTGGATCTTTGATGTTCCCTTAAATTTTGTATGTGACA      | 2040 |
| DB | 1981 | TTATCTTCTCGAGGCGCAGTGTGGATCTTTGATGTTCCCTTAAATTTTGTATGTGACA      | 2040 |
| QY | 2041 | GATAGGAAACCTTGATCCGACCTGAATTCGAAATGCTTTCTAGTGTCTATATACGGGT      | 2100 |
| DB | 2041 | GATAGGAAACCTTGATCCGACCTGAATTCGAAATGCTTTCTAGTGTCTATATACGGGT      | 2100 |
| QY | 2101 | ATGTTGTAATGCTTCTTTGATTTTGTAAAGCATAAAGTGGATATGAGTCACTCTCTGAAT    | 2160 |
| DB | 2101 | ATGTTGTAATGCTTCTTTGATTTTGTAAAGCATAAAGTGGATATGAGTCACTCTCTGAAT    | 2160 |
| QY | 2161 | TTACTGTTTCAGGTGAGGGAATATGAGTCCAGGAGTGTGAGTCCGAAGCCGACATGATTTCTA | 2220 |
| DB | 2161 | TTACTGTTTCAGGTGAGGGAATATGAGTCCAGGAGTGTGAGTCCGAAGCCGACATGATTTCTA | 2220 |
| QY | 2221 | AAAGCTATAGAGAGGAGTCCGAGCGGTGTCGAAAGTGTAGTCTATAGTACGAAGACA       | 2280 |
| DB | 2221 | AAAGCTATAGAGAGGAGTCCGAGCGGTGTCGAAAGTGTAGTCTATAGTACGAAGACA       | 2280 |
| QY | 2281 | GGAGTCTTTTTTTTTTTTTTTTTTTTCAAAACCTTAAGACATATAAGGGATTTTTATTTGT   | 2340 |
| DB | 2281 | GGAGTCTTTTTTTTTTTTTTTTTTTTCAAAACCTTAAGACATATAAGGGATTTTTATTTGT   | 2340 |
| QY | 2341 | AACTTATTTATGAAATTAACCTGACCTTCTAAATGCAATGCAAGTGTGACCTGCTAGATCC   | 2400 |
| DB | 2341 | AACTTATTTATGAAATTAACCTGACCTTCTAAATGCAATGCAAGTGTGACCTGCTAGATCC   | 2400 |
| QY | 2401 | ATTTCCCGCGTTCACGGCCCTTGTAGTCTGCTCGCTATACAGTCTGCTGCTCAAGG        | 2460 |
| DB | 2401 | ATTTCCCGCGTTCACGGCCCTTGTAGTCTGCTCGCTATACAGTCTGCTGCTCAAGG        | 2460 |
| QY | 2461 | TATGAGTGTAGATTTGATCAATTAATATATATATATATATATATATATATATATATAT      | 2520 |
| DB | 2461 | TATGAGTGTAGATTTGATCAATTAATATATATATATATATATATATATATATATATAT      | 2520 |
| QY | 2521 | TCCTGATAGTGAACCTCCTTACTAGCTTTTACTTACAACTTAAGCACTTTTGTGCTT       | 2580 |
| DB | 2521 | TCCTGATAGTGAACCTCCTTACTAGCTTTTACTTACAACTTAAGCACTTTTGTGCTT       | 2580 |
| QY | 2581 | CGGTACACAGTTGAATTTGTTGAGTCTTTTTCCTCATAGTGGACCTAGTCTATGTC        | 2640 |
| DB | 2581 | CGGTACACAGTTGAATTTGTTGAGTCTTTTTCCTCATAGTGGACCTAGTCTATGTC        | 2640 |
| QY | 2641 | CTTGATTTTCTTCTTCTTGTGATGTTTCTATGTCATGCAAACTCCAATATGGGTAAAGG     | 2700 |
| DB | 2641 | CTTGATTTTCTTCTTCTTGTGATGTTTCTATGTCATGCAAACTCCAATATGGGTAAAGG     | 2700 |
| QY | 2701 | TTACCTCCTGTTGGGATTTACAGAGTTCCTTTTCTTCTTACACCTGAATGTTTG          | 2760 |
| DB | 2701 | TTACCTCCTGTTGGGATTTACAGAGTTCCTTTTCTTCTTACACCTGAATGTTTG          | 2760 |
| QY | 2761 | TTTTTTATGTTTGTAGTCTTGACAGAGATGCTCCCATCATATTTAGTCTTTTCCCTTTC     | 2820 |
| DB | 2761 | TTTTTTATGTTTGTAGTCTTGACAGAGATGCTCCCATCATATTTAGTCTTTTCCCTTTC     | 2820 |
| QY | 2821 | TCCTTTGTGTCGTTCTCTCTGATGTTTCTTCTGATTAAGCTTTTACTTTTACTTTT        | 2880 |
| DB | 2821 | TCCTTTGTGTCGTTCTCTCTGATGTTTCTTCTGATTAAGCTTTTACTTTTACTTTT        | 2880 |
| QY | 2881 | TCCAGCGAGGTGAATTTTATACGTATCAACCTCAATATCCGACCTATAATTTAAAGAAC     | 2940 |
| DB | 2881 | TCCAGCGAGGTGAATTTTATACGTATCAACCTCAATATCCGACCTATAATTTAAAGAAC     | 2940 |
| QY | 2941 | ACTTAGCTAGATGTTCACTTTTGAATAATTTATTCGAAAGGGGCGAGATGTTCT          | 3000 |
| DB | 2941 | ACTTAGCTAGATGTTCACTTTTGAATAATTTATTCGAAAGGGGCGAGATGTTCT          | 3000 |
| QY | 3001 | GAGGATGCTCCGATCAATTTACCGTGATCAATTTGTCGGTGTGTTTCTGAT             | 3060 |
| DB | 3001 | GAGGATGCTCCGATCAATTTACCGTGATCAATTTGTCGGTGTGTTTCTGAT             | 3060 |
| QY | 3061 | CTGGCTCACTGCTACGAGCTGGGATTTGATCTAGATGGTCACTAGATGGATCTTTGG       | 3120 |

|    |      |                                                              |      |
|----|------|--------------------------------------------------------------|------|
| DB | 3061 | CTGGCTCACTGCTACGAGCTGGGATTTGGATCTAGATGGTCACTAGATGGATCTTTGG   | 3120 |
| QY | 3121 | ACTGGATTTACAAAGCTGGATTAGCATGAACCTTCTGTTTACGGTCTGGTCTGGT      | 3180 |
| DB | 3121 | ACTGGATTTACAAAGCTGGATTAGCATGAACCTTCTGTTTACGGTCTGGTCTGGT      | 3180 |
| QY | 3181 | CTGCTACTCCGCGCTATCAGCTGTAGGATCTCATCGAAAGTTTGGACTATGATTACT    | 3240 |
| DB | 3181 | CTGCTACTCCGCGCTATCAGCTGTAGGATCTCATCGAAAGTTTGGACTATGATTACT    | 3240 |
| QY | 3241 | CTGATTCCTCAATATATTTATCTTTTGAACAATAGTGGATTTCTGTGTGAGTTCTTTCT  | 3300 |
| DB | 3241 | CTGATTCCTCAATATATTTATCTTTTGAACAATAGTGGATTTCTGTGTGAGTTCTTTCT  | 3300 |
| QY | 3301 | AGGACAGCATTTAAGCTCCCGGACATAGATGGAGATGGTCACTAAATTTCTTTGTTATG  | 3360 |
| DB | 3301 | AGGACAGCATTTAAGCTCCCGGACATAGATGGAGATGGTCACTAAATTTCTTTGTTATG  | 3360 |
| QY | 3361 | CCACCTTACATGGGTTTTCGGTCTTCTGCTGAGGTCCTCAATCAAGATCAAAATCAAGA  | 3420 |
| DB | 3361 | CCACCTTACATGGGTTTTCGGTCTTCTGCTGAGGTCCTCAATCAAGATCAAAATCAAGA  | 3420 |
| QY | 3421 | TCAAGATCAAGATCGAATTTCCAGTTTCCAGTGTGAGTCTAAAGCTGAACCTTCT      | 3480 |
| DB | 3421 | TCAAGATCAAGATCGAATTTCCAGTGTGAGTCTAAAGCTGAACCTTCT             | 3480 |
| QY | 3481 | TTAATTCACAAATCCATGTTTGTTTTAAATACCTGCTCACTTTGTTGTTTCTTCAATCA  | 3540 |
| DB | 3481 | TTAATTCACAAATCCATGTTTGTTTTAAATACCTGCTCACTTTGTTGTTTCTTCAATCA  | 3540 |
| QY | 3541 | CACCACTTAAAGAAATCATGAGACAGACTATAAAATTTGAAGAGTCTGTAGAACACTA   | 3600 |
| DB | 3541 | CACCACTTAAAGAAATCATGAGACAGACTATAAAATTTGAAGAGTCTGTAGAACACTA   | 3600 |
| QY | 3601 | GGTCTCACCAACCTCTGTTGTCACCTAAATCGCTCTCCAGTGTTCAGCAACATAAT     | 3660 |
| DB | 3601 | GGTCTCACCAACCTCTGTTGTCACCTAAATCGCTCTCCAGTGTTCAGCAACATAAT     | 3660 |
| QY | 3661 | CTACCTCTGTCATGTTTATCTTTCTTTAAAGGATTTACATATATGTTTGT           | 3720 |
| DB | 3661 | CTACCTCTGTCATGTTTATCTTTCTTTAAAGGATTTACATATATGTTTGT           | 3720 |
| QY | 3721 | CAGGTGATATCTGGTTGAAAATGAAACTGGCTGTACCCGATTCGCTCAAGC          | 3780 |
| DB | 3721 | CAGGTGATATCTGGTTGAAAATGAAACTGGCTGTACCCGATTCGCTCAAGC          | 3780 |
| QY | 3781 | TTCTCAGGCTCCACTGCTTAATAGAAATTTGATTCGGATTTGGGATTTATATCTGGTCTC | 3840 |
| DB | 3781 | TTCTCAGGCTCCACTGCTTAATAGAAATTTGATTCGGATTTGGGATTTATATCTGGTCTC | 3840 |
| QY | 3841 | TTGATGGGACGACCAATATGTTTCTAGTTTGTAGTGTGACCTGGAAATGGTCTGTT     | 3900 |
| DB | 3841 | TTGATGGGACGACCAATATGTTTCTAGTTTGTAGTGTGACCTGGAAATGGTCTGTT     | 3900 |
| QY | 3901 | ATTGTTGTCATTAATAAGCGGAACTCTGCTCGGCTGCATATAATAAGTTTCATCAGACAT | 3960 |
| DB | 3901 | ATTGTTGTCATTAATAAGCGGAACTCTGCTCGGCTGCATATAATAAGTTTCATCAGACAT | 3960 |
| QY | 3961 | TGTTTGGGTGGTGGAGGTTTTTCCATACATATACATTTACATTAACAATCTGTTGTC    | 4020 |
| DB | 3961 | TGTTTGGGTGGTGGAGGTTTTTCCATACATATACATTTACATTAACAATCTGTTGTC    | 4020 |
| QY | 4021 | TTTTTATGATTTATCTTAAACTAAAC                                   | 4044 |
| DB | 4021 | TTTTTATGATTTATCTTAAACTAAAC                                   | 4044 |

RESULT 2  
 ATH131214  
 LOCUS 5164 bp DNA linear PLN 07-MAY-1999  
 DEFINITION Arabidopsis thaliana srp30 gene, exons 1-12.  
 ACCESSION AJ131214

AJ131214.1 GI:4775269  
SP2/ASF-like splicing modulator; srp30 gene.  
Arabidopsis thaliana (thale cress)  
ORIGIN  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE  
1 Lopato,S., Kalyna,M., Dörner,S., Kobayashi,R., Krainer,A.R. and Barta,A.  
TITLE atSRP30, one of two SP2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes  
JOURNAL Genes Dev. 13 (8), 987-1001 (1999)  
MEDLINE 99234087  
PUBMED 10215626  
REFERENCE  
2 (bases 1 to 5164)  
Barta,A.  
Direct Submission  
TITLE Submitted (02-DEC-1998) Barta A., Institute for Biochemistry,  
JOURNAL University of Vienna, Dr. Bohrgasse 9/3,, A-1030 Vienna, AUSTRIA  
FEATURES  
source  
1. .5164  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/mol\_type="genomic DNA"  
join(1806..2014,2138..2203,2290..2364,2444..2546,  
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,  
3501..3578,4515..4572,4844..5125)  
1806..2014  
/number=1  
1928..4858  
/gene="srp30"  
join(1928..2014,2138..2203,2290..2364,2444..2546,  
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,  
3501..3578,4515..4572,4844..4858)  
/gene="srp30"  
/codon\_start=1  
/product="SP2/ASF-like splicing modulator Srp30"  
/protein\_id="CA842557.1"  
/db\_xref="GI:4775270"  
translations="MSRWNRITVGNLPGDIRKCEVDLFYKGPVIVDILKIPRP  
PGYAFVEFPRADDAIYGRDGYDFDGCRLVEIAHGGRSPSPVDYSSYSASRA  
PSRSYRVLTGLPPSASQDLKDHMRKAGVCFSEVPDRKMGSGVVDYSNYDMK  
YAIKLDATFRNFAFSAYIRVREYSRSVSRSPDDSKYSRSGRSPSCSYSSKSR  
SVSPARISIPRSPRLSRLSVSRSGRSRSGRSRSPSPVPIVG"  
join(1928..2014,2138..2203,2290..2364,2444..2546,  
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,  
3501..3578,4178..4214)  
/gene="srp30"  
/note="alternative exon 10a"  
/codon\_start=1  
/product="SP2/ASF-like splicing modulator Srp30, variant 1"  
/protein\_id="CA842558.1"  
/db\_xref="GI:4775271"  
translations="MSRWNRITVGNLPGDIRKCEVDLFYKGPVIVDILKIPRP  
PGYAFVEFPRADDAIYGRDGYDFDGCRLVEIAHGGRSPSPVDYSSYSASRA  
PSRSYRVLTGLPPSASQDLKDHMRKAGVCFSEVPDRKMGSGVVDYSNYDMK  
YAIKLDATFRNFAFSAYIRVREYSRSVSRSPDDSKYSRSGRSPSCSYSSKSR  
SVSPARISIPRSPRLSRLSRLSVSRSGRSRSGRSRSPSPVPIVG"  
2015..2137  
/gene="srp30"  
/number=1  
2138..2203  
/gene="srp30"  
/number=2  
2204..2289  
/gene="srp30"  
/number=2  
2290..2364  
/gene="srp30"  
/number=3  
2365..2443

/gene="srp30"  
/number=3  
2444..2546  
/gene="srp30"  
/number=4  
2547..2631  
/gene="srp30"  
/number=4  
2632..2678  
/gene="srp30"  
/number=5  
2679..2785  
/gene="srp30"  
/number=5  
2786..2843  
/gene="srp30"  
/number=6  
2844..2967  
/gene="srp30"  
/number=6  
2968..3017  
/gene="srp30"  
/number=7  
3018..3160  
/gene="srp30"  
/number=7  
3161..3217  
/gene="srp30"  
/number=8  
3218..3290  
/gene="srp30"  
/number=8  
3291..3403  
/gene="srp30"  
/number=9  
3404..3500  
/gene="srp30"  
/number=9  
3501..3578  
/gene="srp30"  
/number=10  
3579..4514  
/gene="srp30"  
/number=10  
4178..4572  
/gene="srp30"  
/note="alternative exon"  
/number=11  
4178..4340  
/gene="srp30"  
/note="alternative exon 10a"  
4515..4572  
/gene="srp30"  
/number=11  
4573..4843  
/gene="srp30"  
/number=11  
4844..5125  
/number=12  
BASE COUNT 1501 a 912 c 1038 g 1713 t  
ORIGIN  
Query Match 99.4%; Score 4021; DB 8; Length 5164;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4043; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
QY 1 AGACAAGATGCTTACTTCTTTAAACATGTTTCGAGTTTATGAAAATGATCACCAGCTTC 60  
Db 1121 AGACAAGATGCTTACTTCTTTAAACATGTTTCGAGTTTATGAAAATGATCACCAGCTTC 1180  
QY 61 TAACTATGATCTCTTCCCTGCAAGCGAAGATGATGATTAACGGAATAT 120  
Db 1181 TAACTATGATCTCTTCCCTGCAAGCGAAGATGATGATTAACGGAATAT 1240

|    |      |                                                                    |      |
|----|------|--------------------------------------------------------------------|------|
| Qy | 121  | CGGAACATCACTCAACAAACCAAAAAATTTGGACATCATATCGCAACAAAATTTCAATAGGAA    | 180  |
| Db | 1241 | CGGAACATCACTCAACAAACCAAAAAATTTGGACATCATATCGCAACAAAATTTCAATAGGAA    | 1300 |
| Qy | 181  | AAATACTGAAATTTCCAAAAACAAAAACCAAAACGGAAACAGACAGGAAACTCAGGACTGA      | 240  |
| Db | 1301 | AAATACTGAAATTTCCAAAAACAAAAACCAAAACGGAAACAGACAGGAAACTCAGGACTGA      | 1360 |
| Qy | 241  | GAGAGACCGTGGACCGTGTCAACGGCGGAAATAGTCTTTGGACGGAGTTACTAATCGGCGA      | 300  |
| Db | 1361 | GAGAGACCGTGGACCGTGTCAACGGCGGAAATAGTCTTTGGACGGAGTTACTAATCGGCGA      | 1420 |
| Qy | 301  | ATTGAGATTTGAGAGGTGGTAGTAGAGGAAACCGAGAGAAATGTTTCTCTCAAAAAAATCCC     | 360  |
| Db | 1421 | ATTGAGATTTGAGAGGTGGTAGTAGAGGAAACCGAGAGAAATGTTTCTCTCAAAAAAATCCC     | 1480 |
| Qy | 361  | CAAGTGTTCCGATCTAGTGTCTCTTTGTGTCAAAAACGACAGTGTTTAGGAAACCTAGG        | 420  |
| Db | 1481 | CAAGTGTTCCGATCTAGTGTCTCTTTGTGTCAAAAACGACAGTGTTTAGGAAACCTAGG        | 1540 |
| Qy | 421  | AGAAATGAATGACCCGATGGTCCGAATCCGAATTCGAAATTCGGTTCGAAATTCGTAATACTAA   | 480  |
| Db | 1541 | AGAAATGAATGACCCGATGGTCCGAATTCGAAATTCGGTTCGAAATTCGTAATACTAA         | 1600 |
| Qy | 481  | CATACAAATTTCCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACCAATAT       | 540  |
| Db | 1601 | CATACAAATTTCCGGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACCAATAT       | 1660 |
| Qy | 541  | ATTAAACAGGCCATTAATAACATATGGGCCGATCTTGATCAACTGGGCTATTCATCGTGTGA     | 600  |
| Db | 1661 | ATTAAACAGGCCATTAATAACATATGGGCCGATCTTGATCAACTGGGCTATTCATCGTGTGA     | 1720 |
| Qy | 601  | TACATGGCGCGCACAGGATTAATAACGAGTTCGGTTTTATAAAGGATACTAGTTTCCA         | 660  |
| Db | 1721 | TACATGGCGCGCACAGGATTAATAACGAGTTCGGTTTTATAAAGGATACTAGTTTCCA         | 1780 |
| Qy | 661  | AACGAAACGGTGGTGTCTCTTTCCGAAACAACTAAACGGTTCCTCGAAACATCTTCTTCT       | 720  |
| Db | 1781 | AACGAAACGGTGGTGTCTCTTTCCGAAACAACTAAACGGTTCCTCGAAACATCTTCTTCT       | 1840 |
| Qy | 721  | TCCTCTTTCTCGAAATPATTTTTCCAGTAATCAATTTCTCTCTCTCTCTCTCTCTCTCTCT      | 780  |
| Db | 1841 | TCCTCTTTCTCGAAATPATTTTTCCAGTAATCAATTTCTCTCTCTCTCTCTCTCTCTCTCT      | 1900 |
| Qy | 781  | AACTAATTTTCTGCTCTGAGGTATCAGATGAGTAGCCGATGGAAATCGTAGCTACGTT         | 840  |
| Db | 1901 | AACTAATTTTCTGCTCTGAGGTATCAGATGAGTAGCCGATGGAAATCGTAGCTACGTT         | 1960 |
| Qy | 841  | GGGAATTTGCTTGAGATATTCGCAAGTGTGAGTTGAAGATCTCTTCTCAACAGGTTTGA        | 900  |
| Db | 1961 | GGGAATTTGCTTGAGATATTCGCAAGTGTGAGTTGAAGATCTCTTCTCAACAGGTTTGA        | 2020 |
| Qy | 901  | AAATTTCTCTCTTTCTCTCGATAAAAAATGAAATTCATTATGACTAGTTTGGGTTCAFAAA      | 960  |
| Db | 2021 | AAA- TTCTCTTTCTCTCGATAAAAAATGAAATTCATTATGACTAGTTTGGGTTCAFAAA       | 2079 |
| Qy | 961  | TTTGCAATTTCTGCTTCTGCTGAGACAAATTAATTCGACTCTTTATGTATATTTGTTTCACTA    | 1020 |
| Db | 2080 | TTTGCAATTTCTGCTTCTGCTGAGACAAATTAATTCGACTCTTTATGTATATTTGTTTCACTA    | 2139 |
| Qy | 1021 | TGGACCAATTTGTGGACATGATTTGAGATTCACCGAGACCTTCCTGGTTATGCCTTTGT        | 1080 |
| Db | 2140 | TGGACCAATTTGTGGACATGATTTGAGATTCACCGAGACCTTCCTGGTTATGCCTTTGT        | 2199 |
| Qy | 1081 | CGAGGTATATTGATCAAGTACAAATTTGTTTTTTTTTTCTTTCTTTCTTTGTAATAGTATAGGCTA | 1140 |
| Db | 2200 | CGAGGTATATTGATCAAGTACAAATTTGTTTTTTTTTTCTTTCTTTCTTTGTAATAGTATAGGCTA | 2259 |
| Qy | 1141 | ATGACTAAGATAGTTTGTTATTTGGTGGCAGTTTGAAGATCCTCGTGTAGCAGACGATGCA      | 1200 |
| Db | 2260 | ATGACTAAGATAGTTTGTTATTTGGTGGCAGTTTGAAGATCCTCGTGTAGCAGACGATGCA      | 2319 |

|    |      |                                                                    |      |
|----|------|--------------------------------------------------------------------|------|
| QY | 1201 | ATTATATGACGCGTATGGTTATGATATTTTGATGGGTGTCGACTCCGGGTATAGTAAACGCATG   | 1260 |
| DB | 2320 | ATTATATGACGCGTATGGTTATGATATTTTGATGGGTGTCGACTCCGGGTATAGTAAACGCATG   | 2379 |
| QY | 1261 | ATGAAAGCTAGCTTAATTTTCTGTAATTTCTTGTAAGAGGTATATCTTTGTGCTGATGTTT      | 1320 |
| DB | 2380 | ATGAAAGCTAGCTTAATTTTCTGTAATTTCTTGTAAGAGGTATATCTTTGTGCTGATGTTT      | 2439 |
| QY | 1321 | TTAGGTTGAGATTGCACATGGTGGTGGTAGATTTTACCACATCAGTTGATAGGTACAGCAG      | 1380 |
| DB | 2440 | TTAGGTTGAGATTGCACATGGTGGTGGTAGATTTTACCACATCAGTTGATAGGTACAGCAG      | 2499 |
| QY | 1381 | CAGCTACAGTCGAGCGGTGCACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT         | 1440 |
| DB | 2500 | CAGCTACAGTCGAGCGGTGCACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT         | 2559 |
| QY | 1441 | TCCTCGATTGTTATTTTGGTGTCTGTAATTTTATATATTTGAAACTCATTTTACTAC          | 1500 |
| DB | 2560 | TCCTCGATTGTTATTTGGTGTCTGTAATTTTATATTTGAAACTCATTTTACTAC             | 2619 |
| QY | 1501 | CTAAACATGAGTGCCTTGTGACCGGATTAACCGCTTCTGCTTCGTGGCAGGACCTTAAGG       | 1560 |
| DB | 2620 | CTAAACATGAGTGCCTTGTGACCGGATTAACCGCTTCTGCTTCGTGGCAGGACCTTAAGG       | 2679 |
| QY | 1561 | TAAAGGACACTATATAGTCTTTTTCTGAAATGTTGGTTCCTATATCATATGTTTGGATT        | 1620 |
| DB | 2680 | TAAAGGACACTATATAGTCTTTTTCTGAAATGTTGGTTCCTATATCATATGTTTGGATT        | 2739 |
| QY | 1621 | TATCTCTTTTCTGAAATGATGTTATTTTGCTATTTACGGGTGATTAGGATCAATCGGCA        | 1680 |
| DB | 2740 | TATCTCTTTTCTGAAATGATGTTATTTTGCTATTTACGGGTGATTAGGATCAATCGGCA        | 2799 |
| QY | 1681 | AGCTGGAGATGTCGTCTTCTGAAAGTTTCCCTGACCGTAAAGGTGAGTTGACATTGCA         | 1740 |
| DB | 2800 | AGCTGGAGATGTCGTCTTCTGAAAGTTTCCCTGACCGTAAAGGTGAGTTGACATTGCA         | 2859 |
| QY | 1741 | TAGTTTGGATAAGCTTTTGTGATGATGTTAGTAAATTAGTCTTTGTGAAGAGAAATAG         | 1800 |
| DB | 2860 | TAGTTTGGATAAGCTTTTGTGATGATGTTAGTAAATTAGTCTTTGTGAAGAGAAATAG         | 2919 |
| QY | 1801 | GTGTTAAGCATCTGAACCTGCTAAACTCAATTCAGTATTTCTTTGTAGGCATGCTGGGG        | 1860 |
| DB | 2920 | GTGTTAAGCATCTGAACCTGCTAAACTCAATTCAGTATTTCTTTGTAGGCATGCTGGGG        | 2979 |
| QY | 1861 | TTGTGGATTATAGCAACTATGATGATGAAGTACGCAGTAAAGTTTATCTTTTGCAAC          | 1920 |
| DB | 2980 | TTGTGGATTATAGCAACTATGATGATGAAGTACGCAGTAAAGTTTATCTTTTGCAAC          | 3039 |
| QY | 1921 | GCAAACTTCCCTGGACCTTATGCCCTTAGACTGCTTTTGTTCATAGTATACCGAGCTGAAT      | 1980 |
| DB | 3040 | GCAAACTTCCCTGGACCTTATGCCCTTAGACTGCTTTTGTTCATAGTATACCGAGCTGAAT      | 3099 |
| QY | 1981 | TTATCTTCTCGAGGCGCAGTGTGGATCTTTGATGATGTTCCCTTAAATTTTGTATGTACA       | 2040 |
| DB | 3100 | TTATCTTCTCGAGGCGCAGTGTGGATCTTTGATGATGTTCCCTTAAATTTTGTATGTACA       | 3159 |
| QY | 2041 | GATTAAGAAACTTTGATGCCACTGAAATTCGAAATGCTTCTCTAGTGCCTATATACGGGT       | 2100 |
| DB | 3160 | GATTAAGAAACTTTGATGCCACTGAAATTCGAAATGCTTCTCTAGTGCCTATATACGGGT       | 3219 |
| QY | 2101 | ATGTTGTATTGCTTTCTTTGATTTTGTAAAGCATAGTGGATATGGAGTCATCTCTGAAT        | 2160 |
| DB | 3220 | ATGTTGTATTGCTTTCTTTGATTTTGTAAAGCATAGTGGATATGGAGTCATCTCTGAAT        | 3279 |
| QY | 2161 | TTACTGTTCCAGGTGAGGGAATATGATCGAGGAGTGTGAGTCGAAGCCAGATGATTTCTA       | 2220 |
| DB | 3280 | TTACTGTTCCAGGTGAGGGAATATGATCGAGGAGTGTGAGTCGAAGCCAGATGATTTCTA       | 3339 |
| QY | 2221 | AAAGCTATAGAAGCAGGAGTCGAGCGGTGCTCAAGCTGTAGCTATAGTACGAGACA           | 2280 |
| DB | 3340 | AAAGCTATAGAAGCAGGAGTCGAGCGGTGCTCAAGCTGTAGCTATAGTACGAGACA           | 3399 |
| QY | 2281 | GGAGGCTCTTTTTTTTTTTTTTTTTTTTTTTTTCATAAACTTAAGACATATTAAGGATTTTATCTG | 2340 |

[illegible]

|    |      |                                                                 |      |
|----|------|-----------------------------------------------------------------|------|
| Db | 4480 | CCACATTACATGGGGTTTTTGGTCTTGCTGAGGTCCTCAATCAAGATCAAAATCAAGA      | 4539 |
| Qy | 3421 | TCAAGATCAAGATCGAATTTCTCCAGTTTTCACCTGTGGTAAGTCTAAAGCTGAACCTTCT   | 3480 |
| Db | 4540 | TCGAATCAAGATCGAATTTCTCCAGTTTTCACCTGTGGTAAGTCTAAAGCTGAACCTTCT    | 4599 |
| Qy | 3481 | TTAATTTACAATCCAT - GTGTTTGGTTTAAATACCTCTCACTTTGGTTGTTCTTCAATCA  | 3539 |
| Db | 4600 | TTAATTTACAATCCATGGTGTGTTTAAATACCTCTCACTTTGGTTGTTCTTCAATCA       | 4659 |
| Qy | 3540 | ACACCAACTTAACGAATCATGAGACAGACTATAAAATTTGAAGACTCTGTAGAACGACT     | 3599 |
| Db | 4660 | ACACCAACTTAACGAATCATGAGACAGACTATAAAATTTGAAGACTCTGTAGAACGACT     | 4719 |
| Qy | 3600 | AGGCTCACCAACTCTGTGTGCACATAAAATCGCCTCTCCAAAGTGTTTTCAGCAACATAA    | 3659 |
| Db | 4720 | AGGCTCACCAACTCTGTGTGCACATAAAATCGCCTCTCCAAAGTGTTTCAGCAACATAA     | 4779 |
| Qy | 3660 | CTTACCTCTGTCAATGTGTTATCATTTTCTTCTCTCTTAACGGTATTACATATTATGTTTT   | 3719 |
| Db | 4780 | CTTACCTCTGTCAATGTGTTATCATTTTCTTCTCTCTTAACGGTATTACATATTATGTTTT   | 4839 |
| Qy | 3720 | GCAGGTGATATCGTGTGAAATGAAACATGGCCACTGGCTGTACCCGAATCGTCTCAAG      | 3779 |
| Db | 4840 | GCAGGTGATATCGTGTGAAATGAAACATGGCCACTGGCTGTACCCGAATCGTCTCAAG      | 4899 |
| Qy | 3780 | CTTCTCAGGCTCCACTGCTGTAATAGAAATTTGATTTCCGATTTGGGATTATTATCTGGTCTT | 3839 |
| Db | 4900 | CTTCTCAGGCTCCACTGCTGTAATAGAAATTTGATTTCCGATTTGGGATTATTATCTGGTCTT | 4959 |
| Qy | 3840 | CTTGTATGGGACCAATATGCTTTCTAGTTTGTAGTTGTGAACCTCGAAATGGTCTGT       | 3899 |
| Db | 4960 | CTTGTATGGGACCAATATGCTTTCTAGTTTGTAGTTGTGAACCTCGAAATGGTCTGT       | 5019 |
| Qy | 3900 | TATTGTGTCATTAATAAGCCGGAACCTGTCTCGCTGCATAATAAAGTTTCATCAGACA      | 3959 |
| Db | 5020 | TATTGTGTCATTAATAAGCCGGAACCTGTCTCGCTGCATAATAAAGTTTCATCAGACA      | 5079 |
| Qy | 3960 | TTCTGTGGGTGGTGGAGGTTTTTCCATACATATACATTTTACATTACAACCTCTGGTGT     | 4019 |
| Db | 5080 | TTCTGTGGGTGGTGGAGGTTTTTCCATACATATACATTTTACATTACAACCTCTGGTGT     | 5139 |
| Qy | 4020 | CTTTTATGATTTATCTTAAACTAAAC                                      | 4044 |
| Db | 5140 | CTTTTATGATTTATCTTAAACTAAAC                                      | 5164 |

|            |                                   |           |                    |           |                 |
|------------|-----------------------------------|-----------|--------------------|-----------|-----------------|
| RESULT 3   |                                   |           |                    |           |                 |
| T12M4      |                                   |           |                    |           |                 |
| LOCUS      | T12M4                             | 59261 bp  | DNA                | linear    | PLN 23-JUN-1998 |
| DEFINITION | Arabidopsis thaliana chromosome 1 | BAC T12M4 | sequence, complete |           |                 |
|            |                                   |           |                    | sequence. |                 |

sequence.  
AC003114  
VERSION AC003114.1  
GI:3249094  
KEYWORDS HTG.

SOURCE ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eustos II; Brassicales; Brassicaceae; Arabicoapsis.  
 REFERENCE  
 1 (bases 1 to 59261)

REFERENCE  
AUTHORS  
Vyssotskaia, V.S., Schwartz, J.R., Toriumi, M., Yu, G., Oji, O.,  
Kwan, A., Liu S., Li, J., Araujo, R., Au, M., Brendel, V., Buehler, E.,  
Conway, A.B., Conway, A.R., Dewar, K., Feng, J., Kim, C., Kurtz, D.,  
Li, Y., Palm, C.G., Shinn, P., Sun, H., Davis, R.W., Ecker, J.R.,  
Federspiel, N.A. and Theologos, A.  
1 (bases 1 to 59461)

**TITLE** Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete  
**JOURNAL** Unpublished (1998)

**JOURNAL  
REMARK**      Unpublished (1998)  
This sequence is of BAC T12M4 from *Arabidopsis thaliana* chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 11040 bp because we submit only the unique sequence of the clone. However, in order to



|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |       |                                                                                                                                                                      |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| AKPLFMTLSSTNSDISCLITRTALVGMVLYGVGNMIDAHKVPFEMPERBVPVTHV<br>MIGTLNGLDFKALCKLEKPNRTVVSWTIIIDGVARVDKPKKALILLPSRMVACDAI<br>KPNKITLAILPVMNIGDLKMGCSVHAYVKGKGFVPCDITRVNSLIDAYAGCGCIQS<br>AFKFIPIPNRKNLVSWTIMISAFHIGMKEAVSMFKMERLGLKPNRVTMISVLN<br>ACSHGGLIEBEFFNFVNVYKIPDVKHVGLVDMRLRKRLEBAEKIALEPIE<br>EKAVWMLLAGCSVYDDAELAEVRTRKLMELERSHGGDVVLMSNIFCCTGTFELDAQR<br>FRQMDVRGVAKLPGRHSQVT"<br>25594. .28968<br>/gene="T12M4.8"<br>join(25594. .25687,26293. .26400,26495. .26687,26785. .27043,<br>27217. .27321,27408. .27494,27577. .27634,28142. .28209,<br>28295. .28390,28457. .28531,28693. .28722,28807. .28968)<br>/gene="T12M4.8"<br>/note="Match to calreticulin (AtCRTL) mRNA gb U27698 and<br>DNA gb U66344. ESTs gb T45719, gb T22451, gb H36323 and<br>gb AA042519 come from this gene."<br>/codon_start=1<br>/evidence=not experimental<br>/protein_id="AAC24083.1"<br>/db_xref="GI:3249100"<br>/translations="MAKIPSLVSLILIGLVAISAAVIFERFDDGMENRVKSEWK<br>KDNATGEMKHTAGNSGDANDKGIOTSDRYEVAISABFPFSNKKDKTLVFOFVKH<br>EOKLDCGGYKLLSGDVDOCKFGGTPVYSIMFGPDI CGYSTKKVHAILTYNEANHLI<br>KXVPSSETDOLTHVYFILRPDAYSILIDNVEKQSGLYSDWDLPLPKKIKDPKAKK<br>PEWDQSEYISDPEDKPKDGYDIPKEIPDPSKPEDDDEDEGEWGTAPTIPNPEYM<br>GEWKPKQIKNPNTKKEWAPLIDNPFDKDDPELYVFPKLYGLELWQVKSGLFDNV<br>LI CDDPDYAKKLADETWKLKDNIOQNIITLQAEKAAFDDEAKKNBEESKADAPASDA<br>EDRPDEGGDDSDSESABETKSVDSSEBSEKDAHAHVCYHNI TWRFTHP"<br>30117. .30527<br>/gene="T12M4.9" |       | Query Match 86.4%; Score 3493.2; DB 8; Length 59261;<br>Best Local Similarity 99.4%; Pred. No. 0;<br>Matches 3622; Conservative 0; Mismatches 8; Indels 14; Gaps 11; |       |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 1     | AGACAAAGATGCTTACTTCTTAAACATGTTGAGGTTTATGAAAATGATCACAGCTTC                                                                                                            | 60    |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55624 | AGACAAAGATGCTTACTTCTTAAACATGTTGAGGTTTATGAAAATGATCACAGCTTC                                                                                                            | 55683 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 61    | TAAGTATGATGCTTCTTCCCTCCAGCGAACAGTGGAGATGATGATAACGGAATAT                                                                                                              | 120   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55684 | TAAGTATGATGCTTCTTCCCTCCAGCGAACAGTGGAGATGATGATAACGGAATAT                                                                                                              | 55743 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 121   | CGGAACATCACTCAACAAACCAAAATTTGGACATCATATCGCAACAAATTCATAGGAA                                                                                                           | 180   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55744 | CGGAACATCACTCAACAAACCAAAATTTGGACATCATATCGCAACAAATTCATAGGAA                                                                                                           | 55803 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 181   | AAATACTGAAATTCAAAAACAGAAAAACAAACCGAACAGAGCAGGAACCTCAGGACTGA                                                                                                          | 240   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55804 | AAATACTGAAATTCAAAAACAGAAAAACAAACCGAACAGAGCAGGAACCTCAGGACTGA                                                                                                          | 55863 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 241   | GAGAGACCGTGGAGCGGTGTCACGCGGAAAAATGGTCTTGGACGGAGTTACTAATCGGCGA                                                                                                        | 300   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55864 | GAGAGACCGTGGAGCGGTGTCACGCGGAAAAATGGTCTTGGACGGAGTTACTAATCGGCGA                                                                                                        | 55923 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 301   | ATTGAGATTTGAGAGTGGTAGAGGAACCGAGAGATGTTCTCTCAAAAAATCC                                                                                                                 | 360   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55924 | ATTGAGATTTGAGAGTGGTAGAGGAACCGAGAGATGTTCTCTCAAAAAATCC                                                                                                                 | 55983 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 361   | CAAGTGTTCGGATCTAGTGTCTCTTTGTCAAAAACGACAGTGTTTAGGAAACCTAGG                                                                                                            | 420   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55984 | CAAGTGTTCGGATCTAGTGTCTCTTTGTCAAAAACGACAGTGTTTAGGAAACCTAGG                                                                                                            | 56043 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 421   | AGAATGAATACCCGATGGTCCGAATCCGATTCGAAATTTGGTTTGAATTTGTAATACATA                                                                                                         | 480   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 56044 | AGAATGAATACCCGATGGTCCGAATCCGATTCGAAATTTGGTTTGAATTTGTAATACATA                                                                                                         | 56103 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 481   | CATACAAATATTCGGTTTGAATGATAGAAAAACACATCGATCCGGTTAGAACATAT                                                                                                             | 540   |
| Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 56104 | CATACAAATATTCGGTTTGAATGATAGAAAAACACATTCGATCCGGTTAGAACATAT                                                                                                            | 56163 |
| Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 541   | ATTAAACAGGCCCAATTAACATATGGCGGATCTTGATCAACTGGGCTATTTCATCGTTCA                                                                                                         | 600   |

|    |       |                                                                |       |
|----|-------|----------------------------------------------------------------|-------|
| Db | 56164 | ATTAACAGGCCCATTAATAACATATAGGCCGATCTTGATCAACTGGGCTATTTCATCGTTGA | 56223 |
| Qy | 601   | TACATGCGCGCGCACAGGATTAATAATCCAGTCCGTTTATTAATAAGGATACATAGTTTCCA | 660   |
| Db | 56224 | TACATGCGCGCGCACAGGATTAATAATCCAGTCCGTTTATTAATAAGGATACATAGTTTCCA | 56283 |
| Qy | 661   | AACGAAACCGTGGTGTCTCTCTTCCAGAACAAATCTAACAGCTTTCTCGAACATCTCTCT   | 720   |
| Db | 56284 | AACGAAACCGTGGTGTCTCTCTTCCAGAACAAATCTAACAGCTTTCTCGAACATCTCTCT   | 56343 |
| Qy | 721   | TCTTCTTTCTCGAAATTAATTTTCCAGTAACTCAATTTCTCTCTCTCTCTCTCTCTCTCT   | 780   |
| Db | 56344 | TCTTCTTTCTCGAAATTAATTTTCCAGTAACTCAATTTCTCTCTCTCTCTCTCTCTCTCT   | 56403 |
| Qy | 781   | AACTAATTTCTGCTCTGAGGTATCAGATGATGACGAGTGAATTCGATCAGTACGATCT     | 840   |
| Db | 56404 | AACTAATTTCTGCTCTGAGGTATCAGATGATGACGAGTGAATTCGATCAGTACGATCT     | 56463 |
| Qy | 841   | GGGAATTTGCTCGAGATATTCGCAAGTGTGAGGTGAAGATCTCTCTCTCAAGGTTTGA     | 900   |
| Db | 56464 | GGGAATTTGCTCGAGATATTCGCAAGTGTGAGGTGAAGATCTCTCTCTCAAGGTTTGA     | 56523 |
| Qy | 901   | AAATTTCTCTCTCTCTCTCGATAAATTTGAATTCATATGATATTTGTTTTCAGTA        | 960   |
| Db | 56524 | AAATTTCTCTCTCTCTCTCGATAAATTTGAATTCATATGATATTTGTTTTCAGTA        | 56582 |
| Qy | 961   | TTTGCAATTTCTGCTGCTGAGACAAATTTAAATCGACTCTTATGATATTTGTTTTCAGTA   | 1020  |
| Db | 56583 | TTTGCAATTTCTGCTGCTGAGACAAATTTAAATCGACTCTTATGATATTTGTTTTCAGTA   | 56642 |
| Qy | 1021  | TGGACCAATTTGGACATTTGATTTGAAGATTTCCACGAGACCTCTCTGTTATGCTTTGT    | 1080  |
| Db | 56643 | TGGACCAATTTGGACATTTGATTTGAAGATTTCCACGAGACCTCTCTGTTATGCTTTGT    | 56702 |
| Qy | 1081  | CGAGTATATTTGATCAAGTACAAATTTGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT     | 1140  |
| Db | 56703 | CGAGTATATTTGATCAAGTACAAATTTGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT     | 56762 |
| Qy | 1141  | ATGACTAAGATAGTTGTTTATTTGGTGGCAGTTTCAAGATCTCTGATGACGAGATGCA     | 1200  |
| Db | 56763 | ATGACTAAGATAGTTGTTTATTTGGTGGCAGTTTCAAGATCTCTGATGACGAGATGCA     | 56822 |
| Qy | 1201  | ATTTATGACGCTGATGTTTATTTGATGGGTGTGCGATTTCCGGGTTAGTAAACGCGATG    | 1260  |
| Db | 56823 | ATTTATGACGCTGATGTTTATTTGATGGGTGTGCGATTTCCGGGTTAGTAAACGCGATG    | 56882 |
| Qy | 1261  | ATGAAAGCTAGCTTAATTTCTGTAATTTCTGTTAAAGGTGTTATCTTTGTTGATGTTT     | 1320  |
| Db | 56883 | ATGAAAGCTAGCTTAATTTCTGTAATTTCTGTTAAAGGTGTTATCTTTGTTGATGTTT     | 56942 |
| Qy | 1321  | TTAGTTGAGATTCGACATGTTGTTGTTTCCACCATCAGTTGATAGGTACACAG          | 1380  |
| Db | 56943 | TTAGTTGAGATTCGACATGTTGTTGTTTCCACCATCAGTTGATAGGTACACAG          | 57002 |
| Qy | 1381  | CAGCTACAGTCCGAGCGGTGACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT     | 1440  |
| Db | 57003 | CAGCTACAGTCCGAGCGGTGACCTTCAAGACGCTCTGACTACCGCGTTTGTAGAGTCT     | 57062 |
| Qy | 1441  | TCTCGATTTGTTTATTTGTTGTTTCTGTTAAATTTTATATTTGAAAACTCAITTTTACTAC  | 1500  |
| Db | 57063 | TCTCGATTTGTTTATTTGTTGTTTCTGTTAAATTTTATATTTGAAAACTCAITTTTACTAC  | 57122 |
| Qy | 1501  | CTAAACATGATGCTTGTGACCGGATTCGCGCTTCTGCTTCGTGGCAGGACCTTAAGG      | 1560  |
| Db | 57123 | CTAAACATGATGCTTGTGACCGGATTCGCGCTTCTGCTTCGTGGCAGGACCTTAAGG      | 57182 |
| Qy | 1561  | TAAGGACATATATAGTCTTTTCTCTGATGTTGTTTCTCTATATCATGTTTTGATT        | 1620  |
| Db | 57183 | TAAGGACATATATAGTCTTTTCTCTGATGTTGTTTCTCTATATCATGTTTTGATT        | 57242 |
| Qy | 1621  | TATCTCTTTTCTGAAATGATGTTTATTTGCTTATTTTACCGGGTATAGGATCACATCGCAA  | 1680  |
| Db | 57243 | TATCTCTTTTCTGAAATGATGTTTATTTGCTTATTTTACCGGGTATAGGATCACATCGCAA  | 57302 |









```

/genes="F7G19.1"
/notes="Similar to ATP-dependent Clp protease (gb|D90915).
EST gb|N65461 comes from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70396.1"
/db_xref="GI:2342674"
/translation="MASCICQAMNSLPRSSSPSPHPLSSNSGRRNLKTRYPAPRA
KASAKVIMPNRPDPPLSLTAANSPEKLNRPNVADVPYLIDPSPOLMSSPA
QVSAYVAPNRPRTPPDLPLMDLGRVLYIGMPDDDIIFCQLVPAVTELVAELM
YLQWLPKEPIYIINSTGRDGTGTHVCVGAAGQACLLSAGTGGKRFMPHA
KGSEBFDPIFILPVMQQLVNRDILVELLSKTSNVSIVTAVNMRRPYMDAPKA
KEFGVDRVSLGTGVI"
complement(3700..3771)
/product="tRNA-Glu"
/notes="codon recognized: GAA"
/genes="F7G19.2"
/genes="F7G19.2"
join(3942..4046,4391..4507,4747..4854,4973..5164,
5263..5454,5551..5669,5747..5834,5916..6035,6216..6356,
6436..6513)
/genes="F7G19.2"
/notes="Similar to probable Mg-dependent ATPase
(pir|556671). ESTs gb|T46782.gb|AA04798 come from this
gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70397.1"
/db_xref="GI:2342675"
/translation="MAEDTSFEGDQLASMTTDDIGRASLLANEIRILKEESQRTNLD
LESVKIKENQSKIKNLQPLVNLGNIVLELMSPEDDAEDGANIDLDORKGKV
VLKTSRQIFLPLVGLPDPDLKPGELVGNKDSYLILDTLPSEYDSRVKAMEVDEK
PTDYNIDGLEQIQELVEAIVLPMTHEQKSGKSLGVPKGLVLYPGPTKILMAR
ACAAQNTFPLKAGLQVQMFGDGAKLVDAFLAKESPCIIIFIDEIDAIKTRF
DSVSGDREVRQITMBLLQDGFSSDDRIKVIATNRADILDPALMSGRDLRIEF
PHTEARGLIQHRSKMNADVNFELARSTDDFNGAQLKAVCVGAEGLALRRDA
TEVNHDENEGIIQVQAKKASLYYA"
complement(6693..9149)
/genes="F7G19.3"
complement(join(6693..6802,6892..7391,7471..7544,
7619..7721,7805..7920,8032..8127,8253..8639,8799..9149))
/genes="F7G19.3"
/notes="Strong similarity to Oryza NADPH oxidase
(gp|X93301)."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70398.1"
/db_xref="GI:2342676"
/translation="MUSQKLIPTKDRNPVRPFAMNISYFFLENKRIWLVLTWISICI
TLFTWFLQYKRTVEFVNGYCVTVAKGSAETLKFMALILLPVCNRTITWLTKSL
IGSVFPDDNINFHKVAFGIAGVIGLHAIHLACDFPRLHAKNVEPEPMKKFGDE
RPNYGMFKGTDGWTGMVLMVLYVLAQSWFRNRANLPKSLKELTGFNAFWS
HLFLVTVLLIVHGVFVYLSKEWHKTMVYLAVPLLYAFERLIRAPRGAKAVY
LKAVYPGNVLSLYMSKPKGPKYTSQYIYINCSDVSPLOWHPFSITASGDDYLSVH
IRTLGWTQSLKSLKVCQLPSTQSGLFIADIQANNITFPRLIDGPGYGAQD
YRNDVLLVGLIGATPLISIRIDVNNIKNQSIERGNTQHRNRYATKRAYFYW
TRQGLEWFSVWNEVAYDSGEMIELHNYCTSYVEGDARSALITLMQSLHAKSG
IDIVSTVRVTRFARNRWSVFHVAVNHVQVGVFYCGNCTIIGELKRLAQDPSRK
TITKFEFKENF"
complement(9310..10430)
/genes="F7G19.4"
complement(join(9310..9492,9582..9630,9745..9904,
10061..10430))
/genes="F7G19.4"
/codon_start=1
/evidence=not_experimental
/product="F7G19.4"
/protein_id="AAB70399.1"
/db_xref="GI:2342677"
/translation="MESSSEGETNKISRCATGSDNPDEYVEITLEVDERINTMKA
KATRLRVSGLRUKTWKYSUSFASRRDRSKSGAMFALRGLRPIAKNDVAGWGDEVA
MRFDKLVAGPKSKFGKICIGVSSSEFNFELFALVRGRTGTSSTITKELFFWE
QITGNSFDRRLQITFDMVDKLDGRITGDEVKEIIALSASANKLSKINENVDEVAALI

```

```

MEELDRNLGYIEVRCLLSNGYTCLTSLILKTLNLP"
/genes="F7G19.5"
join(11553..11812,11908..12119,12289..13011,13083..13344,
13419..13929)
/genes="F7G19.5"
/notes="Similar to Arabidopsis luminal binding protein
(gb|D89342)."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70400.1"
/db_xref="GI:2342678"
/translation="MSSLAITEGERQKLTGIVIGLTTYSVGVYHNKHVEIIANDQG
NRITPSVAFTDTERLIGEAKNOAKNPRTIFDPKELIGRKFPDQDVORDIKELPY
KYANKDGPYIQVYKGEKLPSEBLSAMILLTKWETAFLGKKIKDAVITVPAYE
NDARQATKQAGATAGLVNRVINEPTGAALAYGLDKKGESNLIVTDLGGTDFDSI
LTIIDGFEVLSTGSDTHLGEDFDRHMDYFIKLVKYKNDIKSOHKALKLURREC
ELAKRSLNQHVSVEISLFDGVDSEPLTRFEELNMDLFKTKEPVKALKADAG
LKSDIDILVLGGTTRILPKVQOOLKDFQFGKPSKGTNPDEAVYGAAGVQVLSGE
GGESTQNLITLLKTDILLDVAPLSLGIEVGVGVTNLIIPRNTVIPTKKSOVFTYQD
QQTVTNVGEGRSMYKDNRELKGLDTGLPAPRNQYFRGVPQISVTEVFDANGIL
QVKAEDKVAKTSQITITNDKRLTEBIEIREAEFEAEEDKIMEKIDARKNLET
YVYMKSTVADKEXLAKKISDEDEKNEGVLKEALEWLENVNAEKEDYDEKLKEVEL
VCDPVIKSVYKTEGENEDTWMITMSYRIV"
complement(14446..15420)
/genes="F7G19.6"
complement(14446..15420)
/genes="F7G19.6"
/notes="Similar to Glycine SRC2 (gb|AB000130). ESTs
gb|H76969.gb|T21700.gb|ATTS089 come from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70401.1"
/db_xref="GI:1922937"
/translation="MECRSLDLTIIAEDLKDVLQIKQDLYAVVISINGDARTKQTK
VDKCGTKPKWQHOMKLTVDAAARDNRNLTVPEIVADRPAGPKPVGEVSPVKELL
DQNGDSEKTYTAVRLPNSKAKGSLKSEKFEKTYTSGSSGPHAPVPSMDKTMW
QPVAYPPGAPASAYPAPFAGSSGYPGQDHDHKGVDGYGYPQAGTGGYPPP
GAYPQGGYGPYGPQQGYPGPYPQGPYGPQGPYPGPYGPQAGHKGPKQPKX
HGKAGAGMGLGLGILGAGLLGGLLVGEAVSDIADMGDMGDMGDMGDF"
17933..21752
/genes="F7G19.7"
join(17933..18538,18639..18781,18879..19029,19082..19181,
19238..19367,19470..20115,20194..20523,20583..20618,
20661..21032,21125..21194,21284..21369,21470..21577,
21678..21752)
/genes="F7G19.7"
/notes="Similar to Vicia sativa ENBP1 (gb|X95995)."
/codon_start=1
/evidence=not_experimental
/protein_id="AAB70402.1"
/db_xref="GI:2342679"
/translation="MQVNPDETCDSVIRMNQETRSANGIGNGSGSIPGIPDLRC
KRSQKQWRCTMSMADKTYCEKHVIOAKRAANSAPRANQKAKKRSSIGEDTYS
GKMDDFELPVTISIDHNNGLASASKNSGRLEKHNKLSRISFETPMRSFSPRAVD

```

Query Match

14.1%; Score 571; DB 8; Length 110514;

Best Local Similarity 99.2%; Pred. No. 2.1e-115;

Matches 506; Conservative 0; Mismatches 0; Indels 5; Gaps 3;

QY 3435 GAATTCCTCCAGTTTCACCTGGTAAGTCTAAAGCTGAACCTTCTTTAATTCAATCC 3494

Db 1 GAATTCCTCCAGTTTCACCT--GTAAGCTAAAGCTGAACCTTCTTTAATTCAATCC 57

QY 3495 AT-CTGTTTGTTTAAATACCTGCTCACTTTGGTTGTTCTTCAATCAACACCACTTAACG 3553

Db 58 ATGTTGTTTGTTTAAATACCTGCTCACTTTGGTTGTTCTTCAATCAACACCACTTAACG 117

QY 3554 AAATCATGACAGACACTATAAAATTTGAAGAGTCTGTAGAACGACTAGGTCTCACCACCC 3613

Db 118 AAATCATGACAGACACTATAAAATTTGAAGAGTCT-TAGAACGACTAGGTCTCACCACCC 176

QY 3614 TCTGTGTGCATAAAATCGCCTCTCCCAAGTGTTCAGAACATAATCTACCTCTGTCTAT 3673

```

Db      177 TCTGTGTCACAAAAATCGCTCTCCAAAGTGTTTCAGAACATAATCTACCTCTGTCAAT 236
Qy      3674 GTGTATCATTTCTTCTCTTAACGGTATTACATATTATGTTTGCAGGTGATATCTG 3733
Db      237 GTGTATCATTTCTTCTCTTAACGGTATTACATATTATGTTTGCAGGTGATATCTG 296
Qy      3734 GTTGAATAAGAAACTGGCCACTGGCTGATCCCAATCGTCTCAAGTCTTCAGGCTCCA 3793
Db      297 GTTGAATAAGAAACTGGCCACTGGCTGATCCCAATCGTCTCAAGTCTTCAGGCTCCA 356
Qy      3794 CTGCTAATAGAAATTTGATTCGGATTGGGATATTATCTAGTCTTCTTGTATGGACGA 3853
Db      357 CTGCTAATAGAAATTTGATTCGGATTGGGATATTATCTAGTCTTCTTGTATGGACGA 416
Qy      3854 CCAATATGCTCTTCTAGTTTGTAGTTGTGAACCTGGAATTCGTCTGTTATGTCATTA 3913
Db      417 CCAATATGCTCTTCTAGTTTGTAGTTGTGAACCTGGAATTCGTCTGTTATGTCATTA 476
Qy      3914 AAAGCCGAAACTCTGCTCGCTGCATATAAAGTTTCACAGCATTTGTTGGGTGTG 3973
Db      477 AAAGCCGAAACTCTGCTCGCTGCATATAAAGTTTCACAGCATTTGTTGGGTGTG 536
Qy      3974 GTGAGGTTTTCCATACATATACATTTACATTTACAACTACTGTTGTTCTTTTATGATTATC 4033
Db      537 GTGAGGTTTTCCATACATATACATTTACATTTACAACTACTGTTGTTCTTTTATGATTATC 596
Qy      4034 TTAACCTAAAC 4044
Db      597 TTAACCTAAAC 607

RESULT 6
LOCUS   AY050912
DEFINITION Arabidopsis thaliana putative SP2/ASF splicing modulator Srp30
ACCESSION AY050912
VERSION   AY050912.1 GI:15292956
KEYWORDS FLI_CDNA.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 933)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
Goldsmith,A.D., Jiang,P.X., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 933)
Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M.,
Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,
Koesena,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Direct Submission
Submitted (24-JUL-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

```

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesena,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGSC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES

```

source
1..933
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAPL08-12-105 (R11038)"
/note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert."
ecotype: Columbia
<1..933
/gene="At1g09140"
<1..626
/gene="At1g09140"
/codon_start=3
/evidence=experimental
/product="putative SP2/ASF splicing modulator Srp30 protein"
/protein_id="AAK93589.2"
/db_xref="GI:21954085"
/translation="IYGRDGYDFGCLRVEIAHGRFRFSPVDYSSYSASRAPSRSDVHLVTGLPPSASMDLKHVRKAGDVCFSVEFDPKMGVGVVDYSDVDMKYAI RKLDAERFNPSSAYINRVREYSRVSRRSPDDSKSYERSRSGPSCSYSKSRSVS FARSISPRFLSRSLSYSSVSRQSKAKSRTRKSNPSVPVISG"
misc_difference 1
/gene="At1g09140"
/note="compared to genomic sequence"
/misc_difference 148
/gene="At1g09140"
/replace="c"
/misc_difference 579
/gene="At1g09140"
/replace="g"
/note="compared to genomic sequence resulting in an amino acid sequence difference"
/replace="t"
627..933
/gene="At1g09140"
BASE COUNT 248 a 187 c 223 g 275 t
ORIGIN
Query Match 7.7%; Score 311.4; DB 8; Length 933;
Best Local Similarity 96.7%; Pred. No. 5,5e-58;
Matches 318; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
3715 GTTTTCAGGTGATATCTGTTTGAATAATGAAACCTGGCCACTGGCTGTACCCGAATCGTC 3774
603 GTTTCACCTGTGATATCTGTTTGAATAATGAAACCTGGCCACTGGCTGTACCCGAATCGTC 662
3775 TCAAGTTCTCAGGCTCCACTGCTAATAGATTTCGATTCGGATTGGGATTATTATATCTG 3834
663 TCAAGTTCTCAGGCTCCACTGCTAATAGATTTCGATTCGGATTGGGATTATTATATCTG 722
3835 GTCTCTTGTATGGACGACCAATATGCTTTCTAGTTTTAGTTTCTGTAACCTGGAAATGG 3894
723 GTCTCTTGTATGGACGACCAATATGCTTTCTAGTTTTAGTTTCTGTAACCTGGAAATGG 782

```



|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AUTHORS               | Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| TITLE                 | Direct Submission                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| JOURNAL               | Submitted (12-JUN-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| COMMENT               | On Jun 12, 2003 this sequence version replaced gi:30962775.<br>----- Genome Center<br>Center: Department Of Chemistry And Biochemistry<br>The University Of Oklahoma<br>Center code:UOKNOR<br>-----<br>* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.<br>* 1 2525: contig of 2525 bp in length<br>* 2626: gap of unknown length<br>* 2626 4919: contig of 2294 bp in length<br>* 4920 5019: gap of unknown length<br>* 5020 9331: contig of 4312 bp in length<br>* 9332 9431: gap of unknown length<br>* 9432 17633: contig of 8202 bp in length<br>* 17634 33329: contig of 15596 bp in length<br>* 33330 33429: gap of unknown length<br>* 33430 73669: contig of 40240 bp in length<br>* 73670 73769: gap of unknown length<br>* 73770 133656: contig of 59887 bp in length.<br>* Location/Qualifiers<br>1. 133656<br>/organism="Medicago truncatula"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:3880"<br>/clone="mth2-28a1"<br>/clone.lib="Medicago truncatula BAC library H2"<br>BASE COUNT 45886 a 22063 c 21955 g 43142 t 610 others<br>ORIGIN<br>1. 133656<br>/organism="Medicago truncatula"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:3880"<br>/clone="mth2-28a1"<br>/clone.lib="Medicago truncatula BAC library H2"<br>BASE COUNT 45886 a 22063 c 21955 g 43142 t 610 others<br>ORIGIN |
| Db                    | 55030 TGAACCTTGACATGCTGGCGGGGATATTCATCTGTAGACCGATAGTAGTATAG 54971                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Qy                    | 1387 CAGTCGCGCCGCTGACCTTCAAGCGCTCTGACTACCGCGTTGTGAGGCTTCTCGA 1446                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Db                    | 54970 TAGTGTGACCGCTGAGTTTCCAGCATCTGAGTATCTGTGTATATATCTATGTTACT 54911                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Qy                    | 1447 TTGTGTATTTTGGTGTGTGTAAATTTTATATTT 1481                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Db                    | 54910 TTTTAAACAGATTCGTAGTTTGTATTTTGT 54876                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| RESULT 9              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| LOCUS                 | AY150486 838 bp mRNA linear PLN 23-SEP-2002                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| DEFINITION            | Arabidopsis thaliana putative SF2/ASF splicing modulator Srp30 (At1g09140) mRNA, complete cds.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| ACCESSION             | AY150486                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| VERSION               | AY150486.1 GI:23297698                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| KEYWORDS              | FLI CDNA.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| SOURCE                | Arabidopsis thaliana (thale cress)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| ORGANISM              | Arabidopsis thaliana                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| REFERENCE             | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| AUTHORS               | 1 (bases 1 to 838)<br>Yamada,K., Chan,W.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| TITLE                 | Arabidopsis Open Reading Frame (ORF) Clones                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| JOURNAL               | Unpublished                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| REFERENCE             | 2 (bases 1 to 838)<br>Yamada,K., Chan,W.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| AUTHORS               | Direct Submission                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| TITLE                 | Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| JOURNAL               | Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| COMMENT               | Location/Qualifiers<br>1. .838<br>/organism="Arabidopsis thaliana"<br>/mol_type="mRNA"<br>/db_xref="taxon:3702"<br>/clone="C105028"<br>/note="This clone is in pUNI 51.<br>ecotype: Columbia"<br>1. .838<br>/gene="At1g09140"<br>1. .807<br>/gene="At1g09140"<br>/codon_start=1<br>/evidence=experimental<br>/product="putative SF2/ASF splicing modulator Srp30"<br>/protein_id="AA13011.1"<br>/db_xref="GI:23297698"<br>/translation="MSSRNRTIYVGNLPGDIRKEVEDLFYKYGPIVDILKIPPRP PGYAFVEDPRDADAIYGRDGYDFDGRRLRVEIAHGRREFFSVDRYSSYSASRA PSRRDYVLVTGLPPSASQDLKHVKAGDVCSEVFPDRKMGSVVDYVSNYDDMK YAIRKLDATEFNFAFSSAYIRVREYKSVSRSDSDSKSYASRSRSPSCSYSKSR SVSPARSISPSRPLSRSLRSLYSVSQSKRSKRSRSPVSPVLSG"<br>3'UTR<br>808. .838<br>/gene="At1g09140"<br>BASE COUNT 212 a 176 c 219 g 231 t<br>ORIGIN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Query Match           | 2.9%; Score 117; DB 8; Length 838;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Best Local Similarity | 92.5%; Pred. No. 4.5e-15;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

```
Matches 123; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2152 TCCTCGAATTACTGTTCAGGTGAGGGAATATGATGCGAGGAGTGAGTCTGAAGCCAG 2211
Db 524 TCCTAGTGTCTTATACGGGTGAGGGAATATGATGCGAGGAGTGAGTCTGAAGCCAG 583
QY 2212 ATGATCTCTAAAGCTATAGAAGCAGGAGTCTGCGAGCGGTGCTCAAGCTGTAGCTATAGTA 2271
Db 584 ATGATCTCTAAAGCTATAGAAGCAGGAGTCTGCGAGCGGTGCTCAAGCTGTAGCTATAGTA 543
QY 2272 GCAAGAGCAGGAG 2284
Db 644 GCAAGAGCAGGAG 656

RESULT 10
AY128356/c
LOCUS AY128356 939 bp mRNA linear PLN 07-AUG-2002
DEFINITION Arabidopsis thaliana unknown protein (At1g09150) mRNA, complete cds.
ACCESSION AY128356
VERSION AY128356.1 GI:22135953
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 939)
AUTHORS Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Palm,C.J., Bower,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (01-JUL-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PDEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAPL cDNAs: Tripp,M.,
Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M.,
Palm,C.J., Bower,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
Location/Qualifiers
1..939
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAPL04-17-K04 (R15763)"
/note="This clone is in pluescript
ecotype: Columbia"
1..939
/gene="At1g09150"
166..711
/gene="At1g09150"
/codon_start=1
/product="unknown protein"
/protein_id="AA091559.1"

gene
CDS
```

```
/db_xref="GI:22135954"
/translation="MPKFCLEISSONQVKSQRRIOSIOIEYEGLESVMEDLLP
KKIPLIVKCPNHLTLVAVNVPILFCIRDGSPYMPITRLLLHOVENIMKRQVDGATK
FVLGNIMCPGLTSPGVLDDQVEAERPVAIYAEQKHALAIGFTWMSADIKSINK
GIGVDNHYLNDGLWNRERLD"
BASE COUNT 276 a 204 c 194 g 265 t
ORIGIN
Query Match 2.8%; Score 113.8; DB 8; Length 939;
Best Local Similarity 98.3%; Pred. No. 2.3e-14;
Matches 115; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 229 CTCACGGACTGAGAGAGACCGTGCACGGTGTACGGCGGAAATGGTCTTGGACGGAGTT 288
Db 117 CTTCCGACTGAGAGAGACCGTGCACGGTGTACGGCGGAAATGGTCTTGGACGGAGTT 58
QY 289 ACTAATCGCGAATTGAGATTGAGCGGTGTAGTAGGAGAACCGAGAGAAATGTTTC 345
Db 57 ACTAATCGCGAATTGAGATTGAGCGGTGTAGTAGGAGAACCGAGAGAAATGTTTC 1

RESULT 11
AX506504
LOCUS AX506504 762 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1199 from Patent WO0216655.
ACCESSION AX506504
VERSION AX506504.1 GI:23387741
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
AUTHORS Harper,J.P., Kreps,J., Wang,X. and Zhu,T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1199 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
Location/Qualifiers
1..762
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 189 a 161 c 197 g 215 t
ORIGIN
Query Match 2.7%; Score 108.8; DB 6; Length 762;
Best Local Similarity 90.6%; Pred. No. 2.9e-13;
Matches 116; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1300 TGTATCTTTGTGTGATGTTTGTAGTTGAGATTGCACATGGTGTGTAGATTTCACC 1359
Db 204 TGATTTTGTATGGGTGTGACTTCGGGTGAGATTGCACATGGTGTGTAGATTTCACC 263
QY 1360 ATCAGTTGATAGGTACAGCAGCAGTACAGTGCAGCGCGTGCACCTTCAAGACGCTCTGA 1419
Db 264 ATCAGTTGATAGGTACAGCAGCAGTACAGTGCAGCGCGTGCACCTTCAAGACGCTCTGA 323
QY 1420 CTACCGCG 1427
Db 324 CTACCGCG 331

RESULT 12
AC126791/c
LOCUS AC126791 120374 bp DNA linear PLN 24-APR-2003
DEFINITION Medicago truncatula clone mth2-7k2, complete sequence.
ACCESSION AC126791
VERSION AC126791.16 GI:29650264
KEYWORDS HTG.
SOURCE Medicago truncatula (barrel medic)
```

ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.  
REFERENCE 1 (bases 1 to 120374)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Medicago truncatula BAC Clone mth2-7k2  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 120374)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (03-JUL-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 3 (bases 1 to 120374)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (19-MAR-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 4 (bases 1 to 120374)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (09-APR-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 5 (bases 1 to 120374)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (10-APR-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 6 (bases 1 to 120374)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (24-APR-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
COMMENT On Apr 9, 2003 this sequence version replaced gi:29124158.  
----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR  
----- Location/Qualifiers  
source 1..120374  
/organism="Medicago truncatula"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3880"  
/clone="mth2-7k2"  
/clone\_lib="Medicago truncatula BAC library H1"  
/note="This is one of two clones in the same well from  
mth2-7k2"  
BASE COUNT 40119 a 19397 c 19969 g 40869 t  
ORIGIN  
Query Match 2.7%; Score 107.6; DB 8; Length 120374;  
Best Local Similarity 57.1%; Pred. No. 3.9e-13;  
Matches 298; Conservative 0; Mismatches 184; Indels 40; Gaps 4;  
QY 1014 TTCAGTATGGACCAATTGGACATTGATTGAAGATTCCACCGAGACTCTCTGTTATG 1073  
DB 70902 TCCAGTATGGTCTTCGTTGACATTGATTGAAGATCCCTCCGAACCAACGAGTTATG 70843  
QY 1074 CTTTGTGCGAGTAT------ATTGATCAAGTACAAATTGTTTCTTCTTC 1122  
||||| |||||||

DB 70842 CTTTGTGAGGTATGTCTACATTTTCATTTTAAAGTAGCGAAGGGGTAGTTATCACAC 70783  
QY 1123 TTGTAATAGTA-----TAGGCTAATGACTAAGATAGTTTGTATTGTTGGTCAGTTGA 1175  
DB 70782 TTTTTTTAGTTCCCTTGTGTTCTGTCTCTAATGTAATTAACATATATTGCGATTGA 70723  
QY 1176 AGATCCTCGTGTATGCGAGCAGTCAATTTATGACGCTGATGTTATGATTTTGAATGGGTG 1235  
DB 70722 GGTATGCTGTGTGCTCAAGATGCGATTATTATAGAGATGTTATGATTTTGAATGTTA 70663  
QY 1236 TCGATCTCGGTTAGTAAACGCAATGATGAAGCTAGCTTAATTTCTCTAATTTCTTCTTA 1295  
DB 70662 TCGATTACTGGTTAGATAATTTTCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 70603  
QY 1296 AAGTGTATCTTTCTGTGATGTT-----TTTAGGTTGAGATTCACAT 1339  
DB 70602 TCGAATTTTGTCAATGACCTTTTGGTCACTATTTGTTACTGTAGTTGAACTTCACAT 70543  
QY 1340 GGTGTGCTGTAGATTTTCCACCATCAGTTGATGTTACAG-----CAGCAGCTACAGTGG 1393  
DB 70542 GGTGACGGGGAAGTTTCATCATCAGTACAGCAGTACAGTACAGCGGTAGAGTGGT 70483  
QY 1394 AGCCGTGCACTTCAAGACGCTCTGACTACCGCGGTTTGTAGAGCTTCTCGAATGTGTT 1453  
DB 70482 AGCCGTGGAGTTTCCAGGCGATCTGACTATCGTGTATCTATTGTTCTCTAATAATTA 70423  
QY 1454 ATTGTTGTTGTGTAATAATTTTATATTTTGAACATCAATTTT 1495  
DB 70422 AAAATATATTTTATGCTTGTATCATTTGTACAGTAGTTT 70381  
RESULT 13  
AC134522/c 129199 bp DNA linear HTG 29-MAR-2003  
LOCUS Medicago truncatula clone mth2-12g4, WORKING DRAFT SEQUENCE.  
DEFINITION  
AC134522  
VERSION AC134522.6 GI:29366966  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
SOURCE Medicago truncatula  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.  
REFERENCE 1 (bases 1 to 129199)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Medicago truncatula BAC Clone mth2-12g4  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 129199)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 3 (bases 1 to 129199)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (29-MAR-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
COMMENT On Mar 29, 2003 this sequence version replaced gi:29336223.  
----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes

```

* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be replaced.
*      1 129199: contig of 129199 bp in length.

```

|                                                                |                                                 |                                                               |       |
|----------------------------------------------------------------|-------------------------------------------------|---------------------------------------------------------------|-------|
| * FEATURES                                                     |                                                 | 1 129199: contig of 129199 bp in length.                      |       |
| source                                                         | Location/Qualifiers                             |                                                               |       |
|                                                                | 1. 129199                                       |                                                               |       |
|                                                                | /organism="Medicago truncatula"                 |                                                               |       |
|                                                                | /mol_type="genomic DNA"                         |                                                               |       |
|                                                                | /db_xref="taxon:3880"                           |                                                               |       |
|                                                                | /clone="mth2-1294"                              |                                                               |       |
|                                                                | /clone.lib="Medicago truncatula BAC library H2" |                                                               |       |
| BASE COUNT                                                     | 42930 a 20950 c 21534 g 43771 t                 | 14 others                                                     |       |
| ORIGIN                                                         |                                                 |                                                               |       |
| Query Match 2.7%; Score 107.6; DB 2; Length 129199;            |                                                 |                                                               |       |
| Best Local Similarity 57.1%; Pred. No. 3.9e-13;                |                                                 |                                                               |       |
| Matches 298; Conservative 0; Mismatches 184; Indels 40; Gaps 4 |                                                 |                                                               |       |
| QY                                                             | 1014                                            | TTCCAGTATGCACCAATTGGACATTGATTGAAGATTCCACCGAGACCTCGCTGGTTATG   | 1073  |
| Db                                                             | 70902                                           | TGCAGTATGTCCTATCGTTGACATTGATTGAAGATCCCTCCGAAACACCACAGGTTATG   | 70843 |
| QY                                                             | 1074                                            | CCCTTGTTCGAGGAT-----ATTGATCAAGTACAAATTTGTTTTTTCTTCTTC         | 1122  |
| Db                                                             | 70842                                           | CTTTTGTGAGGTATGTCACATTTCTATTTTAAAGATGCGAAGGGTAGTTATCACAC      | 70783 |
| QY                                                             | 1123                                            | TTCTAATAGTA-----TAGGCTAATAGTAAGATAGTTGTTATTTGGTGGCAGTTTGA     | 1175  |
| Db                                                             | 70782                                           | TTTTTTTTAGTTCCTTGTGTTGTTCTGCTCATGTATCTTAAACATTTATATTCAGTTTTGA | 70723 |
| QY                                                             | 1176                                            | AGATCTCGTGATGCAGACGATCAATTTATGGACGTGATGTTATGATTTTGGTGGGTG     | 1235  |
| Db                                                             | 70722                                           | GGATGTCGTGATGTCTCAAGATCGGATTTATTATAGAGATGGTTATGATTTTGATGGTTA  | 70663 |
| QY                                                             | 1236                                            | TCGACTTCGGGTTAGTAAACGCATGATGAAGCTAGCTTAATTTTCTGTAATTTCTTGTA   | 1295  |
| Db                                                             | 70662                                           | TCGATTACTGGTTAGAAATATTTTCCCTTCTCTTCTGCTTCTATGGGACCACTA        | 70603 |
| QY                                                             | 1296                                            | AAGGTGTATCTTTTGTGTGATGT-----TTTTAGTTGAGATTGCACAT              | 1339  |
| Db                                                             | 70602                                           | TCGAAATTTTGTCTATATGACCTTTTGGTCACTATTTGTTACTGTAGGTTGAACTTCACAT | 70543 |
| QY                                                             | 1340                                            | GGTGGTGTAGATTTTTCACATCAGTTGATAGTACAG-----CAGCAGCTACAGTGGG     | 1393  |
| Db                                                             | 70542                                           | GGTGGACGGGGAAGTTTCATCATCAGTAGACCGGATACAGTAGGCATAGCCGTAGAGTGGT | 70483 |
| QY                                                             | 1394                                            | AGCCGCTGCACCTTCAAGACGCTCTGACTACCGGGTTTGTAGAGTCTTCTCGATTGTGTT  | 1453  |
| Db                                                             | 70482                                           | AGCCGTGGAGTTTCCAGGCGATCTGACTATCGTGGTATCTATTGTTTCTTAATAATTA    | 70423 |
| QY                                                             | 1454                                            | ATTTGGTGTGTGTAAATTTTATATTGAAACTCATTTTT                        | 1495  |
| Db                                                             | 70422                                           | AAAATAATTTTTTAGCTTTGTATCATCTGTACAGTAGTATTTTT                  | 70381 |

|            |                                                                                                                                                                                                   |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RESULT 14  |                                                                                                                                                                                                   |
| AC012329   |                                                                                                                                                                                                   |
| LOCUS      | 80367 bp DNA linear PLN 19-JAN-2001                                                                                                                                                               |
| DEFINITION | Arabisidopsis thaliana chromosome 1 BAC TIG12 genomic sequence,<br>complete sequence.                                                                                                             |
| ACCESSION  | AC012329                                                                                                                                                                                          |
| VERSION    | AC012329.5 GI:12324433                                                                                                                                                                            |
| KEYWORDS   | HTG.                                                                                                                                                                                              |
| SOURCE     | Arabidopsis thaliana (thale cress)                                                                                                                                                                |
| ORGANISM   | Arabidopsis thaliana                                                                                                                                                                              |
|            | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;<br>rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsi- |
| REFERENCE  | 1 (bases 1 to 80367)                                                                                                                                                                              |
| AUTHORS    | Lian,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.J.,<br>Wu,D., Maiti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R.,                                                          |

|           |                                                                                                           |
|-----------|-----------------------------------------------------------------------------------------------------------|
| TITLE     | Barnstead, W. E., Bowman, C. L., White, O., Nierman, W. C. and Fraser, C. M.                              |
| JOURNAL   | Arabidopsis thaliana chromosome 1 BAC TIG12 genomic sequence                                              |
| REFERENCE | 2 (bases 1 to 80367)                                                                                      |
| AUTHORS   | Lin, X. and Kaul, S.                                                                                      |
| TITLE     | Direct Submission                                                                                         |
| JOURNAL   | Submitted (23-OCT-1999) The Institute for Genomic Research, 9712                                          |
| REFERENCE | 3 (bases 1 to 80367)                                                                                      |
| AUTHORS   | Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org                                                |
| TITLE     | Town, C. D. and Kaul, S.                                                                                  |
| JOURNAL   | Direct Submission                                                                                         |
| REFERENCE | Submitted (19-JAN-2001) The Institute for Genomic Research, 9712                                          |
| AUTHORS   | Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org                                              |
| COMMENT   | On Jan 19, 2001 this sequence version replaced gi:12280863.<br>Address all correspondence to: at@tigr.org |

BAC clone TIG12 is from *Arabidopsis thaliana* chromosome 1. The orientation of the sequence is from SP6 to T7 end of the BAC clone.

Genes were identified by a combination of several methods: Gene prediction programs including Genscan\* (Chris Burge, <http://CCR-081.mit.edu/GENSCAN.html>), GenemarkHMM (Mark Borodovsky, <http://genemark.biology.gatech.edu/Genemark/>), GlimmerA (a variant of Glimmer, see Mihaela Perte, <http://www.tigr.org/soflab/glimmerm.htm/glimmerm.html>, and genesplicer (Mihaela Perte and Steven Salzberg, <http://genesplicer.tigr.org/>), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arlian Smith, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

```

FEATURES             Location/Qualifiers
     source            1..80367
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /cultivar="Columbia"
                        /db_xref="taxon:3702"
                        /chromosome="1"
                        /clone="TIG12"
                        /complement(1616..1637)
repeat_region        /rpt_family="AT_rich"
                        /complement(1661..1713)
repeat_region        /rpt_family="(TA)n"
                        /complement(3271..3326)
repeat_region        /rpt_family="AT_rich"
                        /complement(4982..5033)
repeat_region        /rpt_family="AT_rich"
                        5803..5823
                        /rpt_family="(A)n"
                        /complement(5873..5870)
                        /gene="TIG12.1"
                        /note="contains Pfam profile: PF00083 sugar (and other)
                        transporter"
mrna                  6578..6943,6097..6199,6278..6493,
                        complement(join<5873..5934,6097..6199,6278..6493,
                        6578..6943,6980..7250,7617..7785,7878..8022,8630..>8780))
                        /gene="TIG12.1"
                        /complement(join(5873)..5934,6097..6199,6278..6493,
                        6578..6943,6980..7250,7617..7785,7878..8022,8630..8780))
                        /gene="TIG12.1"
                        /codon_start=1
                        /product="putative transporter; 8780-5873"
                        /protein_id="AG52147.1"
                        /db_xref="GI:12324434"
CDS                   /translation="MEVPYLVFGYMAAVAALELSTKNKORINTSSFNFKNYLL
                        VFSIMAGDWLGGPVYVYLYLVYGFKGIDIGOLFIAFGCSMLFTQIVIGSLAQGRK

```



```

RACVYICIVILSCITKHSPOYKVLVGVGRILGGIATSLIFAPESWLIHAKNRNPEQ
QWLSLTSKAVFNGVAILGLFGLNLLVDTFSGVPAPFDAACFLAIGMAILIGT
WSENFQSPDSKDLITQFKVAATAIADEXALIGALQSLFEASMTFVFLWTPALSP
NDSEIPGHFVAFPMLASMGSSGLARLMSRLRVENTYMQIVFLVSAASLLPIITS
VLVTPSKVDEGLSISIQILGFCVFESCVFWSIPKORSQYIPEEARSTIMNFF
RVPLNFVCIVLXNVDAFFITIMFGMCSIFLVSILQRLMLVISEKPAEDWSPMKE
RNSVDVLTLL
8543. 8562
/rept_family="CAAT)n"
/complement(9170. .11431)
/gene="TlG12.2"
/notes="similar to hypothetical protein (Yer156cp)
GB:6321004 [Saccharomyces cerevisiae]"
/complement(9170. .9394,9488. .9559,9692. .9811,
10121. .10312,10653. .10751,10822. .10968,11209. .>11431))
/gene="TlG12.2"
/complement(9170. .9394,9488. .9559,9692. .9811,
10121. .10312,10653. .10751,10822. .10968,11209. .11418))
/gene="TlG12.2"
/codon_start=1
/product="unknown protein; 11418-9170"
/protein_id="AAG52181.1"
/db_xref="GI:1232441"
/translation="MFGSRGLCCSRIWNRSLFLKRSNFRASPRKRVGTHNGTHCD
EALACFILRSNRFSDAQIVTRDHOVLEKLDALDVGVDYDQSERVYDHQKGFSEV
FLGNTKLRSAGLVYKYLEILSKELQLEORHPVFLFLAYKNFIEAVDALNG
IHQYDTPQPRVYNNVSLGHRIGLNDLIEPOSSAKDEAFHRAVELASBFLSCV
HFHAKSLWPARIVMECLAKRYDIDSSGEINKLSKQCPWKLFHFESEEMKIDPPIKY
VLQYDRESNWIQAVSVSPERFESKALPLAWRGEKEKLSBESSIPRCVFMHSGF
IGANQTEGALAMARASIMA"
12217. .12837
/gene="TlG12.3"
/notes="predicted by genscan"
<12217. .>12837
/gene="TlG12.3"
12217. .12837
/gene="TlG12.3"
/codon_start=1
/product="hypothetical protein; 12217-12837"
/protein_id="AAG52190.1"
/db_xref="GI:12324450"
/translation="MILPKLSLEHIKTPYEOMAYVHLLMSLLIFITKSSISVVP
DPPTQNLNIRICKDTIDFKCNQITISQVRRQTSIKTIKLAIAKAWINAIQTLDMI
EDSLPKARXKRAEDFADKRLVDAHLNDAKLYLRDYFRFMYQALALVNIS
MCRTSFFHPTPMVYNNMKMLTDIAIVAGKILSPPPPLTKSKSP"
/complement(13030. .13063)
/rept_family="AT rich"
/complement(14136. .15366)
/gene="TlG12.4"
/notes="similar to cysteine proteinase GB:AAA50755 [Ainus
glutinosus]; contains Pfam profile: PF00112 papain family
cysteine protease"
/complement(914136. .14716,14922. .>15366))
/gene="TlG12.4"
/complement(914136. .14716,14922. .15366))
/gene="TlG12.4"
/codon_start=1
/product="putative cysteine proteinase; 15366-14136"
/protein_id="AAG52191.1"
/db_xref="GI:12324451"
/translation="MTSIVFFLLAILLSRSTSGVTSRGLFPEASAVEKHEQWMSRFR
VYDSDSEKTSRFEITNNLKFEVSNMNTKTYTLDVNEFSDLTDEBFKARYTGLVVP
EGMTLITTDSEHTVSRFVYNGETGESMDWIQEGAVSVKHQQCCGCCWAFSAVAV
EGMTKIANGELVLSSEQLLDCSTENNGCGGIMKAPDXIKENOGITTEDNYPYOGA
QCTCSNHLAAATLSGYETVPCNDDEALLKAVSQPVSVLAIEGSGYFIHYSGGIFNG
ECGTQTHAVTIVGVGSEEGIKYWLKNSWGSENGYINRIMEDVDSPOGMGLAS
LAYPVA"
/complement(16903. .16941)
/rept_family="AT rich"
/complement(17275. .17301)
/rept_family="AT rich"
/complement(18455. .20638)
/gene="TlG12.5"

```

```

/notes="contains pfam profile: PF00566 TBC domain (implies
that these domains are GTPase activator proteins of
Rab-like small GTPases)"
complement(914136. .14716,14922. .15366,19970. .20019,
20116. .20251,20481. .>20638))
/gene="TlG12.5"
/complement(914136. .14716,14922. .15366,19970. .20019,
20116. .20251,20481. .20638))
/gene="TlG12.5"
/codon_start=1
/product="putative GTPase activator protein of Rab-like
small GTPases; 20638-18455"
/protein_id="AAG52193.1"
/db_xref="GI:12324453"
/translation="WKALRRITQTTSSSNPSSPLSSPPSSSSSWIQRSAFLFVAS
SSPASCSPRLKSPKSPERKEKRLRPOQWKEFTPDGLRNGGVDLKKVRSRGIEP
SIRLEWFLGLGLYFNSSKEKERVINRRSFFDSLAHRCYKCRKEVELRBOCKR
LOKHNGRFLKNGSETIQEDYMPRVDTUSSCSDEVSSEVYVQFSATPDVNSTYP
SEVSTCVERDDGSSRRITNATISTNSSESSDSSEVYVQFSATPDVNSTYP
ASSSIPRTEEDFTVQRIIRLDVADASEMTSPSQAVISDRACRAAEAVGLKDYN
HLEPKYIQAARLVAVLAVAYLDPDICYGQCSMDLLSPILSVIPDDHEVFWCFVGFPM
KKARHFRLEDVGRRLQNLIVSKIIKSKDSQLYSHLEKLOAEDCFVYVMVVMFRRE
LTDLQTLCLVEVMWADQAIIRAGMKSARSIRQAPPTDVLVYALAAASVLQKRRI
IERVMSDEILRECOQSMAGQDVLWKLDDAHLVVLTHKIEHSFS"
/complement(23036. .23068)
/rept_family="AT rich"
/complement(23851. .25075)
/gene="TlG12.6"
/notes="similar to 6-phosphogluconolactonase GB:CAB57866
[Homo sapiens]"
/complement(914136. .14716,14922. .15366,19970. .20019,
20116. .20251,20481. .20638))
/gene="TlG12.6"
/complement(914136. .14716,14922. .15366,19970. .20019,
20116. .20251,20481. .20638))
/gene="TlG12.6"
/codon_start=1
/product="putative 6-phosphogluconolactonase; 25075-23981"
/protein_id="AAG52194.1"
/db_xref="GI:12324454"
/translation="MAPVKRRVPKTNEMAVELAKYADLSKFKCKRGVFTVLSGG
DLIAWLKMLEAPYIDSIENSKWHLFWDERVCAWDHADSNIKLYADFLSKVPFAE
Query Match 2.6%; Score 105.4; DB 8; Length 80367;
Best Local Similarity 54.0%; Pred. No. 1.2e-12;
Matches 289; Conservative 0; Mismatches 231; Indels 15; Gaps 3;
Qy 717 TTCTTCTCTCTCGAAATATTTTCAGTAAATCAATTTCTCTCTCTAGATTTT 776
Db 53367 TTCTTGTCTATAGCAGTAAATTTTATCTCTTATTTATTCGTTCTCTTTTATTC 53426
Qy 777 CAGGAA--CTAATTTCTCTCTGAGTATCAGATGAGTAGCGATGGAATCGTACGATC 834
Db 53427 TGTGCAAGTTTCCCTTTTAACTTTGCAACAATGAGTGGCGGATTTTCTCGTCAATC 53486
Qy 835 TAGCTTGGGAATTTGCTCGAGATATTCGCAAGTGTGAGTTGAAGATCTCTTTTACAG 894
Db 53487 TAGTTGGTAACTTGCCTCGTACATAGGGAACATGAGATTTGAAGATATCTTTTACAG 53546
Qy 895 GTTTGAAAATTTCTCTCTCTCGATATAATGAAATTCATTTATGACTAGTTGGGTT 954
Db 53547 GTTCAGTTTGTGTTATTTGTTGATTCAGTTAGCGGTCTCTCTTGAATCGTTATCCCTA 53606
Qy 955 CATAAATTTGCAATTCGTCTGCTGAGACAATTTAAATCGACTCTTATGATATTTGTT 1014
Db 53607 TAAGACAGTGAAGATTTCTTG----AGCATCTTTTGTATTCATATTTAATGCTCTGTTT 53661
Qy 1015 TCAGTATGACCAATTTGCGACATTTGAGATTTGAGATTTCCACCGAGACCTCTGTTGATGC 1074
Db 53662 GCAGTATGCCGCAATTTGCGATATTGAATTGAAGGTTCCACCTCGCCCTCCATGTTATG 53721
Qy 1075 CTTTGTGAGGATATTTGATCAAGTAAATTTGTTTTTTTCTTCTCTTCTTGTAAATAGTAT 1134
Db 53722 CTTTGTGAGGTAAGTATTTGCTTACATTACATGTTCTTCCACCACTTTTAAATTTTCGC 53781

```

Qy 1135 AGCTAATGACTAAGATAGTTGGTTATTCGGTGGCAGTTGTAAGATCTTCGTGATGCAGAC 1194

Db 53782 TGAACCAAGCTC-----TGTGAATCTTTGTATGTTTGAACATCTTCGGGATGCTGAA 53833

Qy 1195 GATGCAATTATGACGCTGATGGTTATGATGTTTGGTGTGCGACTTCCGGTTA 1249

Db 53834 GATGCCATCAAGGCCGCTGATGCTATAATTTGGATGGCTGCTGCTTGAAGGTAA 53888

RESULT 15

ATT9C5 104204 bp DNA linear PLN 02-FEB-2000

LOCUS Arabidopsis thaliana DNA chromosome 3, BAC clone T9C5.

DEFINITION Arabidopsis thaliana (thale cress)

ACCESSION AL132964

VERSION AL132964.2 GI:6561941

KEYWORDS Arabidopsis thaliana

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE 1 Rieger, M., Gabel, C., Mueller-Auer, S., Schaefer, M., Zipp, M., Mewes, H.W., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M. Unpublished

AUTHORS EU Arabidopsis sequencing project.

JOURNAL Direct Submission

REFERENCE 2 (bases 1 to 104204)

AUTHORS Submitted (02-FEB-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de

TITLE Coordinator: Marcel Salanoubat and Francis Quetier, Groupement d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cremieux, BP191, 91006 Evry Cedex, France;

JOURNAL http://www.genoscope.cns.fr

COMMENT On Dec 12, 1999 this sequence version replaced gi:6434246. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES

Location/Qualifiers

1..104204

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/variety="Columbia"

/db\_xref="taxon:3702"

/chromosome="3"

misc\_feature 1..2572

/notes="overlap to BAC F2K15, please refer to EMBL:AL132956 for analysis and annotation"

gene 667..4510

/gene="T9C5.10"

CDs join(667..601,984..1227,1309..1398,1498..1588,1689..1755,1247..2416,2495..2593,2734..2802,3011..3149,3314..3515,3928..4055,4140..4239,4344..4510)

/genes="T9C5.10"

/notes="similarity to transcription factor IIIC63 - Homo sapiens, EMBL:AF131324

Contains Prokaryotic membrane lipoprotein lipid attachment site AA491-501"

/codon\_start=1

/product="putative protein"

/protein\_id="CAB62446.1"

/db\_xref="GI:6561942"

/translation="MGIIIEGTISGLTPSKEAFVHVFPYPSSTSRRAIETLGGTGGIT QARETSNKLEIRFPEDYAPALGECQPCSGFLLR:SKQDIKPESQSVLDY:SRDV CIEEAGPVLCAIDIVARLSFFHFGDMADYQHVIP:HADIAQOKKKNWMDVPLTGKSD LMGLEDVVMLLPQFAPK:PDNVALKPPATSGPKKKDDAATONFVEIPKLLKWD FVRSNSHHQWQVARSALFEERPIWTRDSVQORLLDKGLKCTHMLNEFLKRAAYES SGFLPLRWKRGDPRNDSERVYQRMERFVPELRGCDANATNSPDSNDICAFK VFPETGFDVFKS IQEDEFERSKQVQKTELTPSLVGREATKQSDMEDTFKSVRNEY DANVNDGSDNLDDESEBDEELMDAAGNEISLDSHGCILOSIAFSCIAGIKTE KTSICFDSPESPNNVYGPVAVDNGSGRFOIVFERSGILNSIDHDNDDEEREDDD

exon 4344..4510  
/gene="T9C5.10"  
/number=13  
gene 5507..6722  
/gene="T9C5.20"  
CDS join(5507..5537,5770..5861,5962..6042,6163..6260,  
6343..6424,6684..6722)  
/gene="T9C5.20"  
/note="Gene model changed according to ESTs GB:A1997074  
and R65541  
similarity to CGI-141 protein- Homo sapiens,  
EMBL:AF151899"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB62447.1"  
/db\_xref="GI:6561943"  
translation="MMSFENDRKKIGLGTGRVPSFLGIIVFDKGLAMGNILF  
ISGSLIIGKSTQTFPMKQNYKGTISFGVGFVILGWPILOMLETYGFVLPFG  
FWPTLAVFAQKIPVLGIIQQPIYIRSFDFYRGRVPEV"  
exon 5507..5537  
/gene="T9C5.20"  
/number=1  
intron 5538..5769  
/gene="T9C5.20"  
/number=1  
exon 5770..5861  
/gene="T9C5.20"  
/number=2  
intron 5862..5961  
/gene="T9C5.20"  
/number=2  
exon 5962..6042  
/gene="T9C5.20"  
/number=3  
intron 6043..6162  
/gene="T9C5.20"  
/number=3  
exon 6163..6260  
/gene="T9C5.20"  
/number=4  
intron 6261..6342  
/gene="T9C5.20"  
/number=4  
exon 6343..6424  
/gene="T9C5.20"  
/number=5  
intron 6425..6683  
/gene="T9C5.20"  
/number=5  
exon 6684..6722  
/gene="T9C5.20"  
/number=6  
gene 8320..10019  
CDS join(8320..8406,8526..8591,8670..8744,8899..9040,  
9124..9170,9255..9312,9449..9502,9587..9638,9869..10019)  
/gene="T9C5.30"  
/note="strong similarity to PRE-MRNA SPLICING FACTOR SF2 -  
Arabidopsis thaliana, SWISSPROT:SPR1\_ARATH"  
/codon\_start=1  
/product="PRE-MRNA SPLICING FACTOR SF2-like protein"  
/protein\_id="CAB62448.1"  
/db\_xref="GI:6561944"  
translation="MSGFRSRISYVGNLPGDIRHEIEDIFYKYGRIVDIELKVPVPRP  
PCYCFVEFHSRDAEDAIKGRDGYNLDELHNGRGQSSSDRRGGVGGGGGY

Query Match 2.6%; Score 105.4; DB 8; Length 104204;  
Best Local Similarity 54.0%; Pred. No. 1.2e-12;  
Matches 289; Conservative 0; Mismatches 231; Indels 15; Gaps 3;  
QY 717 TTCTTCTCTCTCGAATATTTTCCAGTAATCAATTTCTCTCTCTCTAGATTTTTA 776  
DB 8227 TTCTTGTGCTATAGCAGTAATTTTGAATCTTATTTATTCGTTCTCTTTTATCTCA 8286

QY 777 CAGGAA--CTAATTTTCTGCTCTGAGGTATCAGATGATGAGCCGATGGAATCGTACGATC 834  
DB 8287 TGTGCAAGTTTCCCTTTTAAACCTTTGCAAAACAATGAGTGGCGGATTTTCTCGGCAATC 8346  
QY 835 TACGTTGGGAATTTGCCTGGAGATATTCGCAAGTGTGAGTTGAAGATCTCTTCTACAAAG 894  
DB 8347 TATGTTGGTAATCTCCCGGTGACATTAGGGAACATGAGATTGAAGATATCTTTTACAG 8406  
QY 895 GTTTGAAAAATTTCCCTCTTTTCTCTCGATAAAAAATTTGAATTCATTATGACTAGTTTGGGTT 954  
DB 8407 GTCAAGTTTGGTTATTTGTTTGTATTCAGTTACGGGTCTCTCTTTGAATCGTTATCCCTA 8466  
QY 955 CATAAATTTGCAATTTCTGTCTTGTCTGAGACAATTTAAATTCGACTCTTATGTATATTGTT 1014  
DB 8467 TAAGACAGTGAAGATTCTTG-----AGCATCTTTTGTATTATTAATTAATGCTCTTGT 8521  
QY 1015 TCAGTATGGACCAATTTGTGGACATTGATTGAAGATTCCACGAGACCTCTCGGTTATGC 1074  
DB 8522 GCAGTATGGCCCATTTGTCGATATTGAATTTGAAGTTCCACCTCGGCTCCATGTTATG 8591  
QY 1075 CTTTGTGAGGTATATTGATCAAGTACAAATTTGTTTTTTTCTTCTTGTGTAATAGTAT 1134  
DB 8582 CTTTGTGAGGTAGTGTATTGCTTACATTACATGTTCCCTAACCACTTTTAAATTCGC 8641  
QY 1135 AGGCTAATGACTAAGATAGTTTGTATTGTTGGTGGCAGTTTGAAGATCCCTCGTATGCAGAC 1194  
DB 8642 TGAACCAAGCTC-----TGTTGAATGTTTGTAGTTTGACATTTCTCGGATCTGAA 8693  
QY 1195 GATGCAATTTATGGACGTGATGTTATGATTTGATGGGTGTCGACTTCGGGTTA 1249  
DB 8694 GATGCCATCAGGGCCGCTGATGCTATAATTTGGATGGCTGTCTGCTGAGGATAA 8748

Search completed: January 28, 2004, 21:00:52  
Job time : 14437 secs







US-09-938-842A-3893

Query Match 19.3%; Score 780; DB 10; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3.8e-187;  
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| QY | 28  | GTTGAGGTTTATTGAAAATGATCACAGGTTCTAACTATGGTATCTTCTTCCCTGCAAG   | 87  |
| DB | 1   | GTTGAGGTTTATTGAAAATGATCACAGGTTCTAACTATGGTATCTTCTTCCCTGCAAG   | 60  |
| QY | 88  | CGAACAGTGGAGATGATTGTAACGGAATATCGGAACATCATCTACACCAACCAAAAT    | 147 |
| DB | 61  | CGAACAGTGGAGATGATTGTAACGGAATATCGGAACATCATCTACACCAACCAAAAT    | 120 |
| QY | 148 | TTGGACATCATATCGCAACAAATTCATAGCAAAATATCTGAAATTCGCAAAACAGAAAA  | 207 |
| DB | 121 | TTGGACATCATATCGCAACAAATTCATAGCAAAATATCTGAAATTCGCAAAACAGAAAA  | 180 |
| QY | 208 | CCAAACGGAACAGACGAGGAATCTACGGAATGAGAGAGACCGGTGAGCGGTGTACGGCGG | 267 |
| DB | 181 | CCAAACGGAACAGACGAGGAATCTACGGAATGAGAGAGACCGGTGAGCGGTGTACGGCGG | 240 |
| QY | 269 | AAATATGGTCTTGGACGAGTTACTAATTCGGGAAATTTGAGATTGAGAGTGTGTACTAG  | 327 |
| DB | 241 | AAATATGGTCTTGGACGAGTTACTAATTCGGGAAATTTGAGATTGAGAGTGTGTACTAG  | 300 |
| QY | 328 | GAAACGAGAGATGTTTCTCTCAAAAAAATCCCAAGTGTTCGGATCTAGTGTCTCTTT    | 387 |
| DB | 301 | GAAACGAGAGATGTTTCTCTCAAAAAAATCCCAAGTGTTCGGATCTAGTGTCTCTTT    | 360 |
| QY | 388 | TGTCGAAAAACGACAGTGTCTTAGGAAACCTAGGAGATGAATGACCCGATGTCCGAATC  | 447 |
| DB | 361 | TGTCGAAAAACGACAGTGTCTTAGGAAACCTAGGAGATGAATGACCCGATGTCCGAATC  | 420 |
| QY | 448 | CGATTGCAAAATGGTTTCAATTTGTAATACTAACATACATATTCGGTTCGAATGATAA   | 507 |
| DB | 421 | CGATTGCAAAATGGTTTCAATTTGTAATACTAACATACATATTCGGTTCGAATGATAA   | 480 |
| QY | 508 | GAAAAACACATTCGATCCGTTAGAACATATATTAACAGGCCCATTAACACATATGGG    | 567 |
| DB | 481 | GAAAAACACATTCGATCCGTTAGAACATATATTAACAGGCCCATTAACACATATGGG    | 540 |
| QY | 568 | CGGATCTTGATCAACTGGGCTATTTCATCGTTTGATACATGCGCGGCACAGGATTAATC  | 627 |
| DB | 541 | CGGATCTTGATCAACTGGGCTATTTCATCGTTTGATACATGCGCGGCACAGGATTAATC  | 600 |
| QY | 628 | CAGTTCGGTTTTATAAAGATATCTAGTTCGAAACGAGCGGTGTCTCTCTTCAG        | 687 |
| DB | 601 | CAGTTCGGTTTTATAAAGATATCTAGTTCGAAACGAGCGGTGTCTCTCTTCAG        | 660 |
| QY | 688 | AACATCTAACGTTTCTCGAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT    | 747 |
| DB | 661 | AACATCTAACGTTTCTCGAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT    | 720 |
| QY | 748 | TAATCAATTTCTCTCTCTAGATTTTACAGGAACATAATTTCTGCTCTGAGGTATCAG    | 807 |
| DB | 721 | TAATCAATTTCTCTCTCTAGATTTTACAGGAACATAATTTCTGCTCTGAGGTATCAG    | 780 |

RESULT 3  
US-09-938-842A-3893  
; Sequence 3893, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24

```

; Patent No. US20020160378A1
;
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Znu, Tong
;
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
;
; FILE REFERENCE: S01P1300-3
;
; CURRENT APPLICATION NUMBER: US/09/938,842A
;
; CURRENT FILING DATE: 2001-08-24
;
; PRIOR APPLICATION NUMBER: US 60/227,866
;
; PRIOR FILING DATE: 2000-08-24
;
; PRIOR APPLICATION NUMBER: US 60/264,647
;
; PRIOR FILING DATE: 2001-01-16
;
; PRIOR APPLICATION NUMBER: US 60/300,111
;
; PRIOR FILING DATE: 2001-06-22
;
; NUMBER OF SEQ ID NOS: 5379
;
; SEQ ID NO 1199
;
; LENGTH: 762
;
; TYPE: DNA
;
; ORGANISM: Arabidopsis thaliana
;
; US-09-938-842A-1199

```

|    | Query Match           | 2.7%;                                                           | Score 108.8;      | DB 10;    | Length 762; |
|----|-----------------------|-----------------------------------------------------------------|-------------------|-----------|-------------|
|    | Best Local Similarity | 90.6%;                                                          | Pred.No. 8.9e-17; |           |             |
|    | Matches 116;          | Conservative 0;                                                 | Mismatches 12;    | Indels 0; | Gaps 0;     |
| Qy | 1300                  | TGTTATCTTTCTGTGATGTTTATAGGTTCAGATTGCATGGTGGTGGTGTAGATTTTCACC    | 1359              |           |             |
| Db | 204                   | TGATTTTGTAGGGTGTGCATCTTCGGGTTCAGATTGCACATGGTGGTGGTGTAGATTTTCACC | 263               |           |             |
| Qy | 1360                  | ATCAGTTGTATAGGTPACAGCAGCAGCTACAGTCGAGCCGCTGCACCTTTCAGACGCTCTGA  | 1419              |           |             |
| Db | 264                   | ATCAGTTGTATAGGTPACAGCAGCAGCTACAGTCGAGCCGCTGCACCTTTCAGACGCTCTGA  | 323               |           |             |
| Qy | 1420                  | CTACCGCG                                                        | 1427              |           |             |
| Db | 324                   | CTACCGCG                                                        | 331               |           |             |

RESULT 5  
US-09-938-842A-1199  
; Sequence 1199, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1199  
; LENGTH: 762  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1199

Query Match 2.7%; Score 108.8; DB 12; Length 762;  
Best Local Similarity 90.6%; Pred. No. 8.9e-17;  
Matches 116; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

Db      204  TGATTTTGATGGGTGCGACTTCGGGTGAGATTGCACATGGTGGTGTAGATTTTCACC 263
Oy      1360  ATCAGTTGATAGGTACAGCAGCAGCTACAGTGCAGCGCGTGACCTTCAAGACGCTCTGA 1419
Db      264  ATCAGTTGATAGGTACAGCAGCAGCTACAGTGCAGCGCGTGACCTTCAAGACGCTCTGA 323
Oy      1420  CTACCGCG 1427
Db      324  CTACCGCG 331

RESULT 6
US-10-311-455-384
; Sequence 384, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 384
; LENGTH: 6161
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-384

```

| Query Match           | 1.6%           | Score 65.4                                                            | DB 13    | Length 6161 |
|-----------------------|----------------|-----------------------------------------------------------------------|----------|-------------|
| Best Local Similarity | 47.1%          | Pred. No. 3.6e-05                                                     |          |             |
| Matches 201           | Conservative 0 | Mismatches 226                                                        | Indels 0 | Gaps 0      |
| Qy                    | 2455           | TC AAGGTATCAGGTGTAGATTTCGTATCAATATTATATATATAGTAGTTACCCCTTCATGGAATC    | 2514     |             |
| Db                    | 2733           | TGAAATTTGCTATTATTATATATTTAAAGTGATATAATTTAGGGGTATTTAGTATATGTTTA        | 2792     |             |
| Qy                    | 2515           | ACTTGTTCCTTGCATAGTGAACACTCCTTACACAGCTTTTATTTACTTACAACACTAAGCACTTTTTCG | 2574     |             |
| Db                    | 2793           | TAGTGTATTATGTAGTATGAAATTTAAATTTTGAAGAATATTTTATTTATTTTAAATATAGAGTTT    | 2852     |             |
| Qy                    | 2575           | TTGCTTCCGTACACAGTGTGAATTTGTTGAGTCTTTTTTCCCTCATAGTGGACTAGTCTA          | 2634     |             |
| Db                    | 2853           | TATATTTAAATTAATAGTATTATTTTATTTTGTTTTTTTTAGTTTTTGGTAAATATAAAT          | 2912     |             |
| Qy                    | 2635           | TTGTCACCTTGATTTTCTTCCTTGTGTGATGTTTCTTATGTCATGCAAACTCCAAATATGGG        | 2694     |             |
| Db                    | 2913           | GTGTTTTTGTGTTTTATGGAATTTGTTTTATTTTGAATATTTTATGTAATATAGAGTTTATATA      | 2972     |             |
| Qy                    | 2695           | TAAAGGTACCTCCTCTGTTTGGGATTAACCAAGAGTTCCTTTTTCATTTCTTACACGTGAATG       | 2754     |             |
| Db                    | 2973           | GTGTGTGTGATATTTTCGTTTTTGGGTTATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTT        | 3032     |             |
| Qy                    | 2755           | TGTTTGTTTTTTATGTTTTTGAGTCTCTTGACAGAGATGCTCCCATCATATTTTAGTCCTTTT       | 2814     |             |
| Db                    | 3033           | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT      | 3092     |             |
| Qy                    | 2815           | CCCTTCTCTCTTGTCGTGCTCTCTCTGGATGTTTCCTCTCTGATAAAGCCTTTACTTCTTAA        | 2874     |             |
| Db                    | 3093           | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT      | 3152     |             |
| Qy                    | 2875           | CTTTTTT 2881                                                          |          |             |





|                                                                                                                 |      |                                                              |      |  |  |
|-----------------------------------------------------------------------------------------------------------------|------|--------------------------------------------------------------|------|--|--|
| RESULT 9                                                                                                        |      |                                                              |      |  |  |
| US-10-311-455-1670                                                                                              |      |                                                              |      |  |  |
| ; Sequence 1670, Application US/10311455                                                                        |      |                                                              |      |  |  |
| ; Publication No. US20030143606A1                                                                               |      |                                                              |      |  |  |
| ; GENERAL INFORMATION:                                                                                          |      |                                                              |      |  |  |
| ; APPLICANT: OLEK, Alexander                                                                                    |      |                                                              |      |  |  |
| ; APPLICANT: PIEPENBROCK, Christian                                                                             |      |                                                              |      |  |  |
| ; APPLICANT: BERLIN, Kurt                                                                                       |      |                                                              |      |  |  |
| ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation |      |                                                              |      |  |  |
| ; FILE REFERENCE: 5013.1014                                                                                     |      |                                                              |      |  |  |
| ; CURRENT APPLICATION NUMBER: US/10/311,455                                                                     |      |                                                              |      |  |  |
| ; CURRENT FILING DATE: 2002-12-16                                                                               |      |                                                              |      |  |  |
| ; PRIOR APPLICATION NUMBER: PCT/EP01/07537                                                                      |      |                                                              |      |  |  |
| ; PRIOR FILING DATE: 2001-07-02                                                                                 |      |                                                              |      |  |  |
| ; PRIOR APPLICATION NUMBER: DE 10032529.7                                                                       |      |                                                              |      |  |  |
| ; PRIOR FILING DATE: 2000-06-30                                                                                 |      |                                                              |      |  |  |
| ; PRIOR APPLICATION NUMBER: DE 10043826.1                                                                       |      |                                                              |      |  |  |
| ; PRIOR FILING DATE: 2000-09-01                                                                                 |      |                                                              |      |  |  |
| ; NUMBER OF SEQ ID NOS: 2424                                                                                    |      |                                                              |      |  |  |
| ; SEQ ID NO 1670                                                                                                |      |                                                              |      |  |  |
| ; LENGTH: 6668                                                                                                  |      |                                                              |      |  |  |
| ; TYPE: DNA                                                                                                     |      |                                                              |      |  |  |
| ; ORGANISM: Artificial Sequence                                                                                 |      |                                                              |      |  |  |
| ; FEATURE:                                                                                                      |      |                                                              |      |  |  |
| ; FEATURE:                                                                                                      |      |                                                              |      |  |  |
| ; NAME/KEY: unsure                                                                                              |      |                                                              |      |  |  |
| ; LOCATION: 1936                                                                                                |      |                                                              |      |  |  |
| ; OTHER INFORMATION: n is a o r g or c o r t                                                                    |      |                                                              |      |  |  |
| US-10-311-455-1670                                                                                              |      |                                                              |      |  |  |
|                                                                                                                 |      |                                                              |      |  |  |
| Query Match            1.6%; Score 63.2; DB 13; Length 6668;                                                    |      |                                                              |      |  |  |
| Best Local Similarity   44.1%; Pred. No. 0.00014;                                                               |      |                                                              |      |  |  |
| Matches 263; Conservative   0; Mismatches 333; Indels   0; Gaps   0;                                            |      |                                                              |      |  |  |
| QY                                                                                                              | 2286 | TCATTTTGTGTTTTTTTTTTTCAAAACCTAGACACATAAAGGATTTTAATGTAACTT    | 2345 |  |  |
| DB                                                                                                              | 2731 | TTTTTTTTTGTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT    | 2790 |  |  |
| QY                                                                                                              | 2346 | ATTATGAATTAACTCAGCTCTTAATGCATGCAGTGTCTCACCTCGTAGATCCATTTC    | 2405 |  |  |
| DB                                                                                                              | 2791 | TTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTT     | 2850 |  |  |
| QY                                                                                                              | 2406 | CCGCGCTTACGCGCCCCCTTAGTCGTTCTCGCTATACAGCTCTGTCTCAAAGGTATGA   | 2465 |  |  |
| DB                                                                                                              | 2851 | TTTTTGTTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT     | 2910 |  |  |
| QY                                                                                                              | 2466 | GCTTAGATTGATCATATTATATATGATGATGCCCTTCATGGATCACTTGTCTTG       | 2525 |  |  |
| DB                                                                                                              | 2911 | TTTTTGTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT     | 2970 |  |  |
| QY                                                                                                              | 2526 | CATAGTGAACCTCTACTAGCTTTATCTACAACACTAACGACCTTTGTGTCCTCGTA     | 2585 |  |  |
| DB                                                                                                              | 2971 | TTTTTTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTGTTTTTTTTTTT    | 3030 |  |  |
| QY                                                                                                              | 2586 | CACAGTTGAATTTGTTTGAGTCTTTTTTCCCTCATAGTGGACTAGTCTATTGTCACTTGA | 2645 |  |  |
| DB                                                                                                              | 3031 | TTTTTTTTTTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT   | 3090 |  |  |
| QY                                                                                                              | 2646 | TTTTCTTCCTTTGTGATGTTTTCTATGTCATGCAAACTCCAATATGGGTAAAGGTACC   | 2705 |  |  |
| DB                                                                                                              | 3091 | TTTTTTTTTTGTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT    | 3150 |  |  |
| QY                                                                                                              | 2706 | TCCTTGTTGGGATTACACAGTTCCTTTTCATTTCTTACACGTGAATGTTGTTTTTT     | 2765 |  |  |
| DB                                                                                                              | 3151 | TTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT    | 3210 |  |  |
| QY                                                                                                              | 2766 | TATGTTTGTAGTTCTTGACACAGATGCTCCCATCATATTTAGTCCTTTTCTTCTCTTT   | 2825 |  |  |
| DB                                                                                                              | 3211 | TTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT   | 3270 |  |  |
| QY                                                                                                              | 2826 | GTGTCGTCTCTCTCGGATGTTTCCCTCTCATAGAAGCTTTTACTTCTTAACTTTTTT    | 2881 |  |  |

[illegible]

RESULT 11  
US-10-001-843-45/c  
; Sequence 45, Application US/10001843  
; Publication No. US2002013225A1  
; GENERAL INFORMATION:  
; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Herve  
; APPLICANT: Cafferey, Robert  
; APPLICANT: Sun, Yongming

```

; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2109
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (335)
; OTHER INFORMATION:
; OTHER INFORMATION: Clone ID: 34-LIB3057-015-Q1-K1-A6
US-09-983-965-2109

Query Match      1.5%; Score 61.4; DB 10; Length 529;
Best Local Similarity 46.8%; Pred. No. 7.5e-05;
Matches 191; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 2474 TTGTGATCATATTATATATAGTAGTTACCCCTTCATGGATCAGTCTGTTCTTGCATAGTGA 2533
Db 108 TT 167

QY 2534 ACTCCTTACTAGCTTTATTACTTACAACCTAAGCACGCTTTGTGTGCTCCGTACACAGTTG 2593
Db 168 TTTTGTGCTT 227

QY 2594 AATTGCTTGGAGTCTTTTTTCCCTCATAGTGGAGTACTATGTGACACTGATTTCTTC 2653
Db 228 TTTTGTGCTT 287

QY 2654 CTTTGTGTGATGTTTCTATGTATGTCGAACTCCAATAGGTAAGGTTACCTCCTTGTT 2713
Db 288 TTTTGTGCTT 347

QY 2714 TGGGATTACACAGAGTTCCTTTTCATTCTTACACGTTGAATGTGTTGTTTTTATGTTT 2773
Db 348 TGTGTTT 407

QY 2774 GAGTCTTCGACAGAGATGTCCTCCATCATATTTAGTGCCTTTTCCTTCTCTTGTGTCGTT 2833
Db 408 TTTTGTGCTT 467

QY 2834 CTTCTTCGATGTTTCCTTCTCATAAAGCTTTACTTCTTAACCTTTTTT 2881
Db 468 TTTTGTGCTT 515

RESULT 13
US-10-311-455-1186
; Sequence 1186, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1186
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1186

```



|    |      |                                                               |      |
|----|------|---------------------------------------------------------------|------|
| QY | 2651 | TTCCCTTCTTGATGTTTTCTATGTCATGCAAACTCCAAATATGGGTAAAGGTTACCTCCTT | 2710 |
| Db | 4510 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT  | 4569 |
| QY | 2711 | GTTCGGGATTACCAAGATCCCTTTTCATTCTTACACGTGAATGTTGTTTTTATGT       | 2770 |
| Db | 4570 | TTTTTTTCGTTTCGTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT    | 4629 |
| QY | 2771 | TTTGAGTTCTTGACAGAGATGCTCCCATCATATTTAGTCCTTTTCCCTTCTCTTTGTGTC  | 2830 |
| Db | 4630 | TTTTTGTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT  | 4689 |
| QY | 2831 | GTTCCTTCTGGAGTTTTCCTTCTGATAAAGCTTTACTTCTTAACTTTTTT            | 2881 |
| Db | 4690 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTT    | 4740 |

Search completed: January 28, 2004, 21:21:40  
Job time : 1246 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 16:22:39 ; Search time 983 Seconds

(without alignments)  
11105.318 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaaagatcttacttct.....atgattatcttaactaac 4044

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

```

1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID          | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 4042.4 | 100.0       | 4044   | 21 AAC81899 | A. thaliana SRP30  |
| 2          | 780    | 19.3        | 780    | 24 ABZ16088 | Arabidopsis thalia |
| 3          | 283    | 7.0         | 568    | 21 AAC34694 | Arabidopsis thalia |
| 4          | 281.4  | 7.0         | 1213   | 21 AAC40180 | Arabidopsis thalia |
| 5          | 272.6  | 6.7         | 545    | 25 ABX57250 | Arabidopsis thalia |
| 6          | 222.4  | 5.5         | 925    | 21 AAC49248 | Arabidopsis thalia |
| 7          | 175    | 4.3         | 714    | 21 AAC42865 | Arabidopsis thalia |
| 8          | 143.8  | 3.6         | 1041   | 21 AAC46891 | Arabidopsis thalia |

```

c 9 142.2 3.5 919 21 AAC36164 Arabidopsis thalia
10 117 2.9 289 21 AAC49452 Arabidopsis thalia
11 117 2.9 771 21 AAC49442 Arabidopsis thalia
12 108.8 2.7 762 24 ABZ13394 Arabidopsis thalia
c 13 78.6 1.9 5285 19 AAV69915 Arabidopsis ASR-2-
c 14 67.4 1.7 6644 20 AAX33181 Base sequence of t
c 15 67.4 1.7 7372 20 AAX33182 Base sequence of t
c 16 67.4 1.7 7797 20 AAX33180 Cowpox virus bsr f
c 17 67.4 1.7 7996 20 AAX33184 Base sequence of t
18 65.4 1.6 6161 24 ABL32411 Human immune syste
19 64.8 1.6 14006 24 ABL33958 Human immune syste
20 64.4 1.6 6171 24 ABL32788 Human immune syste
c 22 64.2 1.6 556 24 ABQ36996 Oligonucleotide fo
23 63.2 1.6 6668 24 ABL36997 Human immune syste
c 24 62.4 1.5 875 22 AA195044 Human neuroblastom
25 62.4 1.5 3683 25 ABZ10199 Haematopoietic cel
c 26 62 1.5 1257 24 ABZ78031 Human breast speci
27 61.4 1.5 529 25 ABX52180 Bovine EST associa
28 61.2 1.5 6334 24 ABL33213 Human immune syste
29 59.8 1.5 885 21 AAC42654 Arabidopsis thalia
30 59.6 1.5 6668 24 ABL33696 Human immune syste
31 59.4 1.5 6621 25 ABZ10129 Haematopoietic cel
32 59.4 1.5 6621 25 ABZ10243 Haematopoietic cel
33 59.2 1.5 786 24 ABQ33462 Oligonucleotide fo
c 34 59.2 1.5 786 24 ABQ33463 Oligonucleotide fo
35 59 1.5 6338 24 ABL33126 Human immune syste
36 58.8 1.5 16633 24 ABN79984 Human chemically m
37 58.6 1.4 3683 25 ABZ10053 Haematopoietic cel
38 58.6 1.4 6767 22 AAS46608 Tumour suppressor
39 58.4 1.4 9539 22 AAS45346 Chemically pretrea
40 58.4 1.4 9539 24 ABK28179 DNA transcription
41 58.2 1.4 8079 24 ABL92313 Chemically treated
42 58.2 1.4 16033 24 ABL33404 Human immune syste
43 58.2 1.4 17934 24 ABL33719 Human immune syste
44 58 1.4 9539 22 AAS45347 Chemically pretrea
45 58 1.4 9539 24 ABK28180 DNA transcription

```

#### ALIGNMENTS

```

RESULT 1
AAC81899
ID AAC81899 standard; DNA; 4044 BP.
XX AAC81899
AC AAC81899
DT 23-FEB-2001 (first entry)
XX
XX
DE A. thaliana SRP30 cDNA.
XX
XX SR protein; splice-factor activity; plant; developmental behavior;
KW flowering; crop plant; cereal; bean; rice; fruit; ss.
XX
XX Arabidopsis thaliana.
XX
XX WO200065059-A1.
XX
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-AT00100.
XX
XX 23-APR-1999; 99AT-0000727.
XX
XX (OSTP ) OESTERR FORSCH SEIBERSDORF.
XX
XX Barta A, Lopato S, Kalyana M, Dörner S;
XX WPI; 2000-687349/67.
XX P-PSDB; AAB11414, AAB11419.
XX
XX Novel proteins with splice-factor activity in plants, useful e.g. for

```

PT altering flowering time or development, and the nucleic acid that  
PT encodes it -

Claim 5; Fig 1A; 67pp; German.

This invention describes a novel protein (I) with splice-factor activity in plants (I) modifies the choice of splice sites in many plant pre-mRNAs. (I) (also the nucleic acid that encodes them and related vectors or expression systems) are used: (i) to alter splice patterns in plants, or their parts; (ii) to alter developmental behavior of plants; and/or (iii) to delay flowering, particularly by at least 2% relative to the wild type, especially in crop plants such as cereals, beans, rice and fruit.

Sequence 4044 BP; 1061 A; 728 C; 828 G; 1427 T; 0 other;

|                       |         |               |        |              |
|-----------------------|---------|---------------|--------|--------------|
| Query Match           | 100.0%; | Score 4042.4; | DB 21; | Length 4044; |
| Best Local Similarity | 100.0%; | Pred. No. 0;  |        |              |

|    |     |                                                                 |     |
|----|-----|-----------------------------------------------------------------|-----|
| QY | 1   | AGACAAAGATCGTTACTTCTTAAACATGTTTCGAGGTTTATTGAAATGATCACCAGCTTC    | 60  |
| DB | 1   | AGACAAAGATCGTTACTTCTTAAACATGTTTCGAGGTTTATTGAAATGATCACCAGCTTC    | 60  |
| QY | 61  | TAACTATGGTATCTTCTCCCTGCAACGGCAACAGTGGAGATGATTGTAACGGAAATAT      | 120 |
| DB | 61  | TAACTATGGTATCTTCTCCCTGCAACGGCAACAGTGGAGATGATTGTAACGGAAATAT      | 120 |
| QY | 121 | CGGAACATCTCACTCAACAAACCAAAATTTGGACATCATATCGCAACAATTTCAATAGGAA   | 180 |
| DB | 121 | CGGAACATCTCACTCAACAAACCAAAATTTGGACATCATATCGCAACAATTTCAATAGGAA   | 180 |
| QY | 181 | AAATACTGAAATTTCCAAACAGAAAAACCAAAACGGAACAGAGCAGGAACCTCAGCACTGA   | 240 |
| DB | 181 | AAATACTGAAATTTCCAAACAGAAAAACCAAAACGGAACAGAGCAGGAACCTCAGCACTGA   | 240 |
| QY | 241 | GAGACACCGTGGACGGTGCTACCGCGGAATAATGGTCTTTGGAACGAGATTTACTAATCGGGA | 300 |
| DB | 241 | GAGACACCGTGGACGGTGCTACCGCGGAATAATGGTCTTTGGAACGAGATTTACTAATCGGGA | 300 |
| QY | 301 | ATTGAGATTTGAGAGGTGGTAGTAGAGAACCGAGAGATGTTTCTCTCAAAAAAATCCC      | 360 |
| DB | 301 | ATTGAGATTTGAGAGGTGGTAGTAGAGAACCGAGAGATGTTTCTCTCAAAAAAATCCC      | 360 |
| QY | 361 | CAAGTGTTCCGATCTAGTGCTCTTTTGTGCCAAAACGACAGTGTTTAGGAAACCTAGG      | 420 |
| DB | 361 | CAAGTGTTCCGATCTAGTGCTCTTTTGTGCCAAAACGACAGTGTTTAGGAAACCTAGG      | 420 |
| QY | 421 | AGAAATGAATGACCCGATGGTCCGAATCCGATTCGAAATCGTTTCGAAATGTAATAACTAA   | 480 |
| DB | 421 | AGAAATGAATGACCCGATGGTCCGAATCCGATTCGAAATCGTTTCGAAATGTAATAACTAA   | 480 |
| QY | 481 | CATACAAATATTCGGGTGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACAAATAT | 540 |
| DB | 481 | CATACAAATATTCGGGTGTTTGAATGATAAGAAAAACACATTCGATCCGGTTAGAACAAATAT | 540 |
| QY | 541 | ATTACAGGCCCATTTAAACAATATGGGCCGATCTTGATCAACTGGGCTATTCAATCGTTGA   | 600 |
| DB | 541 | ATTACAGGCCCATTTAAACAATATGGGCCGATCTTGATCAACTGGGCTATTCAATCGTTGA   | 600 |
| QY | 601 | TACATCGGCCCGCACAGATTTAAAATCCAGTTCGGTTTTTATAAAGGATACAGTTCCCA     | 660 |
| DB | 601 | TACATCGGCCCGCACAGATTTAAAATCCAGTTCGGTTTTTATAAAGGATACAGTTCCCA     | 660 |
| QY | 661 | AACGAACGGTGCTGCTCCCTTCAGAACCAATCTAACGCTTCTCGAACATCTCTCTCT       | 720 |
| DB | 661 | AACGAACGGTGCTGCTCCCTTCAGAACCAATCTAACGCTTCTCGAACATCTCTCTCTCT     | 720 |
| QY | 721 | TCTTCTTCTCGAAATATTTTTCAGTAATCAATTTCTTCTCTCTAGATTTTACAGG         | 780 |
| DB | 721 | TCTTCTTCTCGAAATATTTTTCAGTAATCAATTTCTTCTCTCTAGATTTTACAGG         | 780 |
| QY | 781 | AACTAAATTTTCTGCTCTGAGGTATCAGATGAGTAGCCCAATCGTACATCTACGTT        | 840 |





[illegible]

PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0134449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 18-JUL-1999; 99US-0144086.  
PR 18-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.

|          |                                                                |
|----------|----------------------------------------------------------------|
| RESULT 4 |                                                                |
| AAC40180 |                                                                |
| ID       | AAC40180 standard; DNA; 1213 BP.                               |
| XX       |                                                                |
| XX       |                                                                |
| AC       | AAC40180;                                                      |
| XX       |                                                                |
| XX       |                                                                |
| DT       | 17-OCT-2000 (first entry)                                      |
| XX       |                                                                |
| XX       |                                                                |
| DE       | Arabidopsis thaliana DNA fragment SEQ ID NO: 27352.            |
| XX       |                                                                |
| KW       | Hybridisation assay; genetic mapping; gene expression control; |
| KW       | protein identification; signal transduction pathway;           |
| KW       | metabolic pathway; promoter; termination sequence; ss.         |
| KW       |                                                                |





```
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 09-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match

5.5%; Score 222.4; DB 21; Length 925;



```
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.3%; Score 175; DB 21; Length 714;
Best Local Similarity 94.8%; Pred. No. 4.2e-32;
Matches 181; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2152 TCTGTAATTTACTGTTCAGGTGAGGGAATATGATCGAGGAGTGTGAGTCGAAGCCAG 2211
Db 524 TCTCTAGTGCTTATATACGGGTGAGGGAATATGATCGAGGAGTGTGAGTCGAAGCCAG 583

Qy 2212 ATGATTTCTAAAGCTATAGAACGAGAGTCGAGCGGTGTCGAAGCTGTAGTATAGTA 2271
Db 584 ATGATTTCTAAAGCTATAGAACGAGAGTCGAGCGGTGTCGAAGCTGTAGTATAGTA 643

Qy 2272 GCAAGAGCAGGAGTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATAAACCTAAGACATATAAGGGAT 2331
Db 644 GCAAGAGCAGGAGTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATAAACCTAAGACATATAAGGGAT 703

Qy 2332 TTTTATTGTAA 2342
Db 704 TTTTATTGTAA 714

RESULT 8
AAC46891/c
ID AAC46891 standard; DNA; 1041 BP.
XX AAC46891;
AC AAC46891;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 51785.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
```



PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140635.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145114.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.

```

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.6%; Score 143.8; DB 21; Length 1041;
Best Local Similarity 98.6%; Pred. No. 1.6e-24;
Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 229 CTCACGACTGAGAGACCGTGCGGTGTCACGGCGGAAATGCTCTTGACCGAGTT 288
Db 150 CTTCCGACTGAGAGACCGTGCGGTGTCACGGCGGAAATGCTCTTGACCGAGTT 91
QY 289 ACTAATCGGCAATTGAGATTTGAGAGGTGGTAGTAGAGAACCGAGAGAAATGTTCTCT 348
Db 90 ACTAATCGGCAATTGAGATTTGAGAGGTGGTAGTAGAGAACCGAGAGAAATGTTCTCT 31

QY 349 CAAAAAAATCCCAAGTGTTCGATC 375
Db 30 CAAAAAAATCCCAAGTGTTCGATC 4

RESULT 9
AAC36164/c
ID AAC36164 standard; DNA; 919 BP.
XX
AC AAC36164;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 12774.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0140931.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

```

|                                                                          |               |
|--------------------------------------------------------------------------|---------------|
| PR 27-JUL-1999;                                                          | 99US-0145913. |
| PR 27-JUL-1999;                                                          | 99US-0145918. |
| PR 27-JUL-1999;                                                          | 99US-0145919. |
| PR 28-JUL-1999;                                                          | 99US-0145951. |
| PR 02-AUG-1999;                                                          | 99US-0146386. |
| PR 02-AUG-1999;                                                          | 99US-0146388. |
| PR 02-AUG-1999;                                                          | 99US-0146389. |
| PR 03-AUG-1999;                                                          | 99US-0147038. |
| PR 04-AUG-1999;                                                          | 99US-0147204. |
| PR 04-AUG-1999;                                                          | 99US-0147302. |
| PR 05-AUG-1999;                                                          | 99US-0147192. |
| PR 05-AUG-1999;                                                          | 99US-0147260. |
| PR 06-AUG-1999;                                                          | 99US-0147303. |
| PR 06-AUG-1999;                                                          | 99US-0147416. |
| PR 09-AUG-1999;                                                          | 99US-0147493. |
| PR 09-AUG-1999;                                                          | 99US-0147935. |
| PR 10-AUG-1999;                                                          | 99US-0148171. |
| PR 11-AUG-1999;                                                          | 99US-0148319. |
| PR 12-AUG-1999;                                                          | 99US-0148341. |
| PR 13-AUG-1999;                                                          | 99US-0148565. |
| PR 13-AUG-1999;                                                          | 99US-0148684. |
| PR 16-AUG-1999;                                                          | 99US-0149368. |
| PR 17-AUG-1999;                                                          | 99US-0149175. |
| PR 18-AUG-1999;                                                          | 99US-0149426. |
| PR 20-AUG-1999;                                                          | 99US-0149722. |
| PR 20-AUG-1999;                                                          | 99US-0149723. |
| PR 20-AUG-1999;                                                          | 99US-0149929. |
| PR 23-AUG-1999;                                                          | 99US-0149902. |
| PR 23-AUG-1999;                                                          | 99US-0149930. |
| PR 25-AUG-1999;                                                          | 99US-0150566. |
| PR 26-AUG-1999;                                                          | 99US-0150884. |
| PR 27-AUG-1999;                                                          | 99US-0151065. |
| PR 27-AUG-1999;                                                          | 99US-0151066. |
| PR 27-AUG-1999;                                                          | 99US-0151080. |
| PR 30-AUG-1999;                                                          | 99US-0151303. |
| PR 31-AUG-1999;                                                          | 99US-0151438. |
| PR 01-SEP-1999;                                                          | 99US-0151930. |
| PR 07-SEP-1999;                                                          | 99US-0152363. |
| PR 10-SEP-1999;                                                          | 99US-0153070. |
| PR 13-SEP-1999;                                                          | 99US-0153758. |
| PR 15-SEP-1999;                                                          | 99US-0154018. |
| PR 16-SEP-1999;                                                          | 99US-0154039. |
| PR 20-SEP-1999;                                                          | 99US-0154779. |
| PR 22-SEP-1999;                                                          | 99US-0155139. |
| PR 23-SEP-1999;                                                          | 99US-0155486. |
| PR 24-SEP-1999;                                                          | 99US-0155659. |
| PR 28-SEP-1999;                                                          | 99US-0156458. |
| PR 29-SEP-1999;                                                          | 99US-0156596. |
| PR 04-OCT-1999;                                                          | 99US-0157117. |
| PR 05-OCT-1999;                                                          | 99US-0157753. |
| PR 06-OCT-1999;                                                          | 99US-0157865. |
| PR 07-OCT-1999;                                                          | 99US-0158029. |
| PR 08-OCT-1999;                                                          | 99US-0158232. |
| PR 12-OCT-1999;                                                          | 99US-0158369. |
| PR 13-OCT-1999;                                                          | 99US-0159293. |
| PR 13-OCT-1999;                                                          | 99US-0159294. |
| PR 13-OCT-1999;                                                          | 99US-0159295. |
| PR 14-OCT-1999;                                                          | 99US-0159329. |
| PR 14-OCT-1999;                                                          | 99US-0159330. |
| PR 14-OCT-1999;                                                          | 99US-0159331. |
| PR 14-OCT-1999;                                                          | 99US-0159637. |
| PR 14-OCT-1999;                                                          | 99US-0159638. |
| PR 18-OCT-1999;                                                          | 99US-0159584. |
| PR 21-OCT-1999;                                                          | 99US-0160741. |
| PR 21-OCT-1999;                                                          | 99US-0160767. |
| PR 21-OCT-1999;                                                          | 99US-0160770. |
| PR 21-OCT-1999;                                                          | 99US-0160814. |
| PR 21-OCT-1999;                                                          | 99US-0160815. |
| PR 22-OCT-1999;                                                          | 99US-0160980. |
| PR 22-OCT-1999;                                                          | 99US-0160981. |
| PR 22-OCT-1999;                                                          | 99US-0160989. |
| PR 25-OCT-1999;                                                          | 99US-0161404. |
| PR 25-OCT-1999;                                                          | 99US-0161405. |
| PR 25-OCT-1999;                                                          | 99US-0161406. |
| PR 26-OCT-1999;                                                          | 99US-0161359. |
| PR 26-OCT-1999;                                                          | 99US-0161360. |
| PR 26-OCT-1999;                                                          | 99US-0161361. |
| PR 28-OCT-1999;                                                          | 99US-0161920. |
| PR 28-OCT-1999;                                                          | 99US-0161992. |
| PR 28-OCT-1999;                                                          | 99US-0161993. |
| PR 29-OCT-1999;                                                          | 99US-0162142. |
| Query Match 3.5%; Score 142.2; DB 21; Length 919;                        |               |
| Best Local Similarity 98.0%; Pred. No. 3.8e-24;                          |               |
| Matches 144; Conservative 0; Mismatches 3; Indels 0; Gaps 0;             |               |
| Qy 229 CTCACGGACTGAGAGAGACCGTGGACGGTGTCAACGGCGGAAAATGCTCTTGGACGGAGTT 288 |               |
| Db 147 CTTCCGACTGAGAGAGACCGTGGACGGTGTCAACGGCGGAGAATGCTCTTGGACGGAGTT 88   |               |
| Qy 289 ACTAATCGGCGAATTGAGATTTGAGAGGTGGTAGAGACCGGACCGAGAGATGTTTCTCT 348   |               |
| Db 87 ACTAATCGGCGAATTGAGATTTGAGAGGTGGTAGAGACCGGACCGAGAGATGTTTCTCT 28     |               |
| Qy 349 CAAAAAATCCCCCAAGTGTTCGGATC 375                                    |               |
| Db 27 CAAAAAATCCCCCAAGTGTTCGGATC 1                                       |               |
| RESULT 10                                                                |               |
| AAC49252                                                                 |               |
| ID AAC49252 standard; DNA; 289 BP.                                       |               |
| XX AAC49252;                                                             |               |
| AC AAC49252;                                                             |               |
| DT 18-OCT-2000 (first entry)                                             |               |
| XX Arabidopsis thaliana DNA fragment SEQ ID NO: 60485.                   |               |
| DE Arabidopsis thaliana                                                  |               |
| XX Hybridisation assay; genetic mapping; gene expression control;        |               |
| KW protein identification; signal transduction pathway;                  |               |
| KW metabolic pathway; promoter; termination sequence; ss.                |               |
| XX Arabidopsis thaliana.                                                 |               |
| OS Arabidopsis thaliana.                                                 |               |
| XX EP1033405-A2.                                                         |               |
| PN 06-SEP-2000.                                                          |               |
| PD 06-SEP-2000.                                                          |               |
| XX 25-FEB-2000; 2000EP-0301439.                                          |               |
| XX 25-FEB-1999;                                                          | 99US-0121825. |
| PR 05-MAR-1999;                                                          | 99US-0123180. |
| PR 03-MAR-1999;                                                          | 99US-0123548. |
| PR 23-MAR-1999;                                                          | 99US-0125788. |
| PR 25-MAR-1999;                                                          | 99US-0126264. |
| PR 29-MAR-1999;                                                          | 99US-0126785. |
| PR 01-APR-1999;                                                          | 99US-0127462. |
| PR 06-APR-1999;                                                          | 99US-0128234. |
| PR 08-APR-1999;                                                          | 99US-0128714. |
| PR 16-APR-1999;                                                          | 99US-0129845. |
| PR 19-APR-1999;                                                          | 99US-0130077. |
| PR 21-APR-1999;                                                          | 99US-0130449. |
| PR 23-APR-1999;                                                          | 99US-0130510. |
| PR 23-APR-1999;                                                          | 99US-0130891. |
| PR 28-APR-1999;                                                          | 99US-0131449. |
| PR 30-APR-1999;                                                          | 99US-0132048. |
| PR 04-MAY-1999;                                                          | 99US-0132407. |
| PR 05-MAY-1999;                                                          | 99US-0132484. |
| PR 06-MAY-1999;                                                          | 99US-0132485. |
| PR 06-MAY-1999;                                                          | 99US-0132486. |
| PR 07-MAY-1999;                                                          | 99US-0132487. |
| PR 11-MAY-1999;                                                          | 99US-0132863. |
| PR 14-MAY-1999;                                                          | 99US-0134256. |
| PR 14-MAY-1999;                                                          | 99US-0134218. |

|    |              |               |
|----|--------------|---------------|
| PR | 14-MAY-1999, | 99US-0134219, |
| PR | 14-MAY-1999, | 99US-0134221, |
| PR | 14-MAY-1999, | 99US-0134231, |
| PR | 14-MAY-1999, | 99US-0134370, |
| PR | 18-MAY-1999, | 99US-0134768, |
| PR | 19-MAY-1999, | 99US-0134941, |
| PR | 20-MAY-1999, | 99US-0135124, |
| PR | 20-MAY-1999, | 99US-0135133, |
| PR | 21-MAY-1999, | 99US-0135629, |
| PR | 24-MAY-1999, | 99US-0136021, |
| PR | 27-MAY-1999, | 99US-0136392, |
| PR | 28-MAY-1999, | 99US-0136782, |
| PR | 01-JUN-1999, | 99US-0137222, |
| PR | 03-JUN-1999, | 99US-0137528, |
| PR | 04-JUN-1999, | 99US-0137502, |
| PR | 07-JUN-1999, | 99US-0137724, |
| PR | 08-JUN-1999, | 99US-0138094, |
| PR | 10-JUN-1999, | 99US-0138540, |
| PR | 10-JUN-1999, | 99US-0138847, |
| PR | 14-JUN-1999, | 99US-0139115, |
| PR | 16-JUN-1999, | 99US-0139145, |
| PR | 18-JUN-1999, | 99US-0139457, |
| PR | 18-JUN-1999, | 99US-0139458, |
| PR | 18-JUN-1999, | 99US-0139459, |
| PR | 17-JUN-1999, | 99US-0139492, |
| PR | 16-JUN-1999, | 99US-0139454, |
| PR | 18-JUN-1999, | 99US-0139455, |
| PR | 18-JUN-1999, | 99US-0139456, |
| PR | 18-JUN-1999, | 99US-0139457, |
| PR | 18-JUN-1999, | 99US-0139462, |
| PR | 18-JUN-1999, | 99US-0139463, |
| PR | 18-JUN-1999, | 99US-0139750, |
| PR | 18-JUN-1999, | 99US-0139763, |
| PR | 21-JUN-1999, | 99US-0139817, |
| PR | 22-JUN-1999, | 99US-0139899, |
| PR | 23-JUN-1999, | 99US-0140353, |
| PR | 23-JUN-1999, | 99US-0140354, |
| PR | 24-JUN-1999, | 99US-0140695, |
| PR | 28-JUN-1999, | 99US-0140823, |
| PR | 28-JUN-1999, | 99US-0140931, |
| PR | 30-JUN-1999, | 99US-0141287, |
| PR | 01-JUL-1999, | 99US-0141842, |
| PR | 01-JUL-1999, | 99US-0142154, |
| PR | 02-JUL-1999, | 99US-0142055, |
| PR | 06-JUL-1999, | 99US-0142390, |
| PR | 08-JUL-1999, | 99US-0142803, |
| PR | 09-JUL-1999, | 99US-0142920, |
| PR | 12-JUL-1999, | 99US-0142977, |
| PR | 13-JUL-1999, | 99US-0143542, |
| PR | 14-JUL-1999, | 99US-0143624, |
| PR | 15-JUL-1999, | 99US-0144005, |
| PR | 16-JUL-1999, | 99US-0144085, |
| PR | 16-JUL-1999, | 99US-0144086, |
| PR | 19-JUL-1999, | 99US-0144325, |
| PR | 19-JUL-1999, | 99US-0144331, |
| PR | 19-JUL-1999, | 99US-0144332, |
| PR | 19-JUL-1999, | 99US-0144333, |
| PR | 19-JUL-1999, | 99US-0144334, |
| PR | 21-JUL-1999, | 99US-0145086, |
| PR | 21-JUL-1999, | 99US-0145088, |
| PR | 20-JUL-1999, | 99US-0144352, |
| PR | 20-JUL-1999, | 99US-0144632, |
| PR | 20-JUL-1999, | 99US-0144884, |
| PR | 21-JUL-1999, | 99US-0144814, |
| PR | 21-JUL-1999, | 99US-0145086, |
| PR | 21-JUL-1999, | 99US-0145088, |
| PR | 22-JUL-1999, | 99US-0145085, |
| PR | 22-JUL-1999, | 99US-0145087, |
| PR | 22-JUL-1999, | 99US-0145089, |
| PR | 22-JUL-1999, | 99US-0145192, |
| PR | 23-JUL-1999, | 99US-0145145, |
| PR | 23-JUL-1999, | 99US-0145218, |
| PR | 23-JUL-1999, | 99US-0145224, |

|    |              |               |
|----|--------------|---------------|
| PR | 26-JUL-1999; | 99US-0145276; |
| PR | 27-JUL-1999; | 99US-0145913; |
| PR | 27-JUL-1999; | 99US-0145918; |
| PR | 27-JUL-1999; | 99US-0145918; |
| PR | 27-JUL-1999; | 99US-0145919; |
| PR | 28-JUL-1999; | 99US-0145919; |
| PR | 28-JUL-1999; | 99US-0145919; |
| PR | 28-JUL-1999; | 99US-0145919; |
| PR | 02-AUG-1999; | 99US-0145886; |
| PR | 02-AUG-1999; | 99US-0146388; |
| PR | 02-AUG-1999; | 99US-0146388; |
| PR | 02-AUG-1999; | 99US-0146389; |
| PR | 03-AUG-1999; | 99US-0147038; |
| PR | 04-AUG-1999; | 99US-0147204; |
| PR | 04-AUG-1999; | 99US-0147302; |
| PR | 05-AUG-1999; | 99US-0147192; |
| PR | 05-AUG-1999; | 99US-0147260; |
| PR | 06-AUG-1999; | 99US-0147303; |
| PR | 06-AUG-1999; | 99US-0147303; |
| PR | 06-AUG-1999; | 99US-0147416; |
| PR | 09-AUG-1999; | 99US-0147493; |
| PR | 09-AUG-1999; | 99US-0147935; |
| PR | 09-AUG-1999; | 99US-0147935; |
| PR | 10-AUG-1999; | 99US-0148171; |
| PR | 11-AUG-1999; | 99US-0148311; |
| PR | 11-AUG-1999; | 99US-0148341; |
| PR | 13-AUG-1999; | 99US-0148565; |
| PR | 13-AUG-1999; | 99US-0148565; |
| PR | 16-AUG-1999; | 99US-0148684; |
| PR | 16-AUG-1999; | 99US-0148684; |
| PR | 17-AUG-1999; | 99US-0149375; |
| PR | 18-AUG-1999; | 99US-0149426; |
| PR | 20-AUG-1999; | 99US-0149722; |
| PR | 20-AUG-1999; | 99US-0149722; |
| PR | 20-AUG-1999; | 99US-0149929; |
| PR | 20-AUG-1999; | 99US-0149929; |
| PR | 23-AUG-1999; | 99US-0149902; |
| PR | 23-AUG-1999; | 99US-0149930; |
| PR | 25-AUG-1999; | 99US-0150566; |
| PR | 26-AUG-1999; | 99US-0150884; |
| PR | 26-AUG-1999; | 99US-0150884; |
| PR | 27-AUG-1999; | 99US-0151065; |
| PR | 27-AUG-1999; | 99US-0151066; |
| PR | 27-AUG-1999; | 99US-0151080; |
| PR | 27-AUG-1999; | 99US-0151303; |
| PR | 30-AUG-1999; | 99US-0151438; |
| PR | 31-AUG-1999; | 99US-0151930; |
| PR | 01-SEP-1999; | 99US-0151930; |
| PR | 07-SEP-1999; | 99US-0152363; |
| PR | 10-SEP-1999; | 99US-0153070; |
| PR | 13-SEP-1999; | 99US-0153758; |
| PR | 15-SEP-1999; | 99US-0154018; |
| PR | 16-SEP-1999; | 99US-0154039; |
| PR | 16-SEP-1999; | 99US-0154039; |
| PR | 20-SEP-1999; | 99US-0154779; |
| PR | 22-SEP-1999; | 99US-0155139; |
| PR | 23-SEP-1999; | 99US-0155486; |
| PR | 23-SEP-1999; | 99US-0155486; |
| PR | 24-SEP-1999; | 99US-0156659; |
| PR | 28-SEP-1999; | 99US-0156458; |
| PR | 29-SEP-1999; | 99US-0156596; |
| PR | 05-OCT-1999; | 99US-0157117; |
| PR | 05-OCT-1999; | 99US-0157753; |
| PR | 06-OCT-1999; | 99US-0157865; |
| PR | 07-OCT-1999; | 99US-0158029; |
| PR | 08-OCT-1999; | 99US-0158232; |
| PR | 12-OCT-1999; | 99US-0158369; |
| PR | 13-OCT-1999; | 99US-0159293; |
| PR | 13-OCT-1999; | 99US-0159294; |
| PR | 13-OCT-1999; | 99US-0159295; |
| PR | 14-OCT-1999; | 99US-0159329; |
| PR | 14-OCT-1999; | 99US-0159330; |
| PR | 14-OCT-1999; | 99US-0159331; |
| PR | 14-OCT-1999; | 99US-0159637; |
| PR | 14-OCT-1999; | 99US-0159638; |
| PR | 18-OCT-1999; | 99US-0159584; |
| PR | 21-OCT-1999; | 99US-0160741; |
| PR | 21-OCT-1999; | 99US-0160767; |
| PR | 21-OCT-1999; | 99US-0160768; |
| PR | 21-OCT-1999; | 99US-0160770; |
| PR | 21-OCT-1999; | 99US-0160814; |
| PR | 21-OCT-1999; | 99US-0160815; |
| PR | 22-OCT-1999; | 99US-0160980; |
| PR | 22-OCT-1999; | 99US-0160981; |
| PR | 22-OCT-1999; | 99US-0160989; |

```
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      2.98; Score 117; DB 21; Length 289;
Best Local Similarity 89.48; Pred. No. 2.9e-18;
Matches 126; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2152 TCTCTGAATTTACTGTTTCAGGTGAGGGAATATGATGTCGAGSAGTGTGAGTCGAAGCCGAG 2211
Db 18 TCTCTAGTCTTATATACGGGTGAGGGAATATGATGTCGAGSAGTGTGAGTCGAAGCCGAG 77

Qy 2212 ATGATTCTAAAGCTATAGAAGCAGGAGTCGGAGCCGTGTCGCAAGCTGTAGCTATAGTA 2271
Db 78 ATGATTCTAAAGCTATAGAAGCAGGAGTCGGAGCCGTGTCGCAAGCTGTAGCTATAGTA 137

Qy 2272 GCAAGAGCAGGAGCTCTTTT 2292
Db 138 GCAAGAGCAGGAGCAGTGTGT 158

RESULT 11
AAC49242
ID AAC49242 standard; DNA, 771 BP.
AC AAC49242;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60449.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
```

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                   |               |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|---------------|
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 22-OCT-1999;                                                      | 99US-0160989. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 25-OCT-1999;                                                      | 99US-0161404. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 25-OCT-1999;                                                      | 99US-0161405. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 25-OCT-1999;                                                      | 99US-0161406. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 26-OCT-1999;                                                      | 99US-0161359. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 26-OCT-1999;                                                      | 99US-0161360. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 26-OCT-1999;                                                      | 99US-0161361. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 28-OCT-1999;                                                      | 99US-0161920. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 28-OCT-1999;                                                      | 99US-0161992. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 28-OCT-1999;                                                      | 99US-0161993. |
| PR                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 29-OCT-1999;                                                      | 99US-0162142. |
| Query-Match            2.9%; Score 117; DB 21; Length 771;<br>Best Local Similarity 92.5%; Pred.No.4.3e-18;<br>Matches 123; Conservative 0; Mismatches 10; Indels 0; Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                   |               |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 2152 TCCTCGAATTACTTTCAGCTGAGGGATATGATCGAGGAGTGTCGAGGCCCCAG        | 2211          |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 524 TCCTAGTGCTTATATACGGGTGAGGAATATGATCGAGGAGTGTCGAGGCCCCAG        | 583           |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 2212 ATGATTCTAAAAGCTATAGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTA | 2271          |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 584 ATGATTCTAAAAGCTATAGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTA  | 643           |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 2272 GCAAGAGCAGGAG                                                | 2284          |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 644 GCAAGAGCAGGAG                                                 | 656           |
| RESULT 12<br>ABZ13394<br>ID ABZ13394 standard; DNA; 762 BP.<br>XX AC ABZ13394;<br>XX DT 21-JAN-2003 (first entry)<br>XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1199.<br>XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.<br>XX OS Arabidopsis thaliana.<br>XX PN WO200216685-A2.<br>XX PD 28-FEB-2002.<br>XX PF 24-AUG-2001; 2001WO-US26685.<br>XX PR 24-AUG-2000; 2000US-227866P.<br>XX PR 26-JAN-2001; 2001US-264647P.<br>XX PR 22-JUN-2001; 2001US-300111P.<br>XX (SCRI ) SCRIPPS RES INST.<br>XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.<br>XX PI Harper JF, Kreps J, Wang X, Zhu T;<br>XX WPI; 2002-304127/34.<br>XX DR Identifying a stress condition to which a plant cell has been exposed<br>XX PT and producing plants with increased tolerance to these abiotic stresses<br>XX PT -<br>XX PS Claim 144; SEQ ID NO 1199; 577pp + Sequence Listing; English.<br>XX CC The invention relates to identifying a stress condition to which a plant<br>XX CC cell has been exposed, comprising:<br>XX CC (a) contacting nucleic acid representative of expressed polynucleotides<br>XX CC in the plant cell with an array or probes representative of the plant<br>XX CC cell genome; and<br>XX CC (b) detecting a profile of expressed polynucleotides in the plant cell<br>XX CC characteristic of a stress response. The method is useful in the plant<br>XX CC cell |                                                                   |               |

production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence, as that of an *Arabidopsis thaliana* stress regulated gene (AB212196-AB217574) used in methods of the invention.

Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office.

| QA | SQ   | Sequence               | 762 BP;                 | 189 A;              | 161 C;         | 197 G;      | 215 T; | 0 other; |
|----|------|------------------------|-------------------------|---------------------|----------------|-------------|--------|----------|
|    |      | Query Match            | 2.7%;                   | Score 108.8;        | DB 24;         | Length 762; |        |          |
|    |      | Best Local Similarity  | 90.6%;                  | Pred. No. 4.1e-16;  |                |             |        |          |
|    |      | Matches 116;           | Conservative            | 0;                  | Mismatches 12; | Indels      | 0;     | Gaps     |
| Qy | 1300 | TGTTATCTTCTGTGATGCTTTT | TAGGTTGACATTGCACATGGT   | CGTGGTCTAGATTTTCACC | 1359           |             |        |          |
| Db | 204  | TGATTTTGATGGGTGTCAC    | TTCCGGTTGAGATTGCACATGGT | CGTGGTCTAGATTTTCACC | 263            |             |        |          |
| Qy | 1360 | ATCAGTTGATAGGTACAGCAGC | AGCTACAGTGCAGCGCGGTGCAC | CTTTCACAGACGCTCTGA  | 1419           |             |        |          |
| Db | 264  | ATCAGTTGATAGGTACAGCAGC | AGCTACAGTGCAGCGCGGTGCAC | CTTTCACAGACGCTCTGA  | 323            |             |        |          |
| Qy | 1420 | CTACCGCG               | 1427                    |                     |                |             |        |          |
| Db | 324  | CTACCGCG               | 331                     |                     |                |             |        |          |

RESULT 13  
AAV69915/C  
ID AAV69915 standard; DNA; 5285 BP.

|    |                                                     |
|----|-----------------------------------------------------|
| XX | AAV69915;                                           |
| AC |                                                     |
| XX | 01-MAR-1999 (first entry)                           |
| DT |                                                     |
| XX | Arabidopsis ASR-2-ORF3 gene region.                 |
| XX |                                                     |
| DE |                                                     |
| XX | Promoter; transgenic plant; ASR-2 gene; vector; ds. |
| XX |                                                     |
| KW |                                                     |
| XX | Arabidopsis thaliana ecotype RDL.                   |
| XX |                                                     |
| OS |                                                     |

| XX | Key      | Location/Qualifiers     |
|----|----------|-------------------------|
| FH | CDS      | 945..3694               |
| FT |          | /*tag= a                |
| FT |          | /label= ASR-2           |
| FT | promoter | complement (3691..4230) |
| FT |          | /*tag= b                |
| FT |          | /note= "Claim 1"        |
| FT | CDS      | complement (4217..4917) |
| FT |          | /*tag= c                |
| FT |          | /label= ORF3            |

WO3844781-A1.  
 15-OCT-1998.  
 03-APR-1998; 98WO-US06761.  
 04-APR-1997; 97US-0042926.  
 (PURD ) PURDUE RES FOUND.

xx  
PI  
xx  
Hodges TK, Lyznik LA;  
xx  
WPI; 1998-609898/51.

PT Novel promoter sequences for expressing genes in eukaryotic cells -  
PT by construction of expression vectors comprising the promoter  
PT sequence useful for high expression of plant proteins  
XX  
PS Example 1; Page 11-14; 23pp; English.

CC This is the nucleotide sequence of the ASR-2 and ORF3 gene region

of *Arabidopsis thaliana*. A genomic library of *A. thaliana* ecotype RDL was screened with a BglI fragment of the rice anther-specific CDNA clone RTS-1. A genomic fragment was isolated and identified as containing sequences homologous to the human splice factor, ASF/SF2 and the *Arabidopsis* SRI gene, and was designated ASR-2. The ASR-2 gene was found to be expressed in all plant parts investigated. The genomic fragment also included ORF3 on the opposite strand relative to ASR-2. A 530 bp region located between the 2 genes was shown to function as a dual promoter. The claimed promoter (see AAU69913) is in the orientation that naturally expresses the ORF3 gene. It is capable of driving strong expression of linked exogenous genes. The claimed promoter can be incorporated into expression vectors to drive high level protein expression in transgenic plants. The exogenous gene may encode products that provide e.g. herbicide, insect or fungal resistance, or which act as growth regulators, or which encode pharmaceutical or polymer components. The promoter is equally, or more, effective than the 35S cauliflower mosaic virus promoter.

XX  
SQ Sequence 5285 BP; 1702 A; 1063 C; 1066 G; 1454 T; 0 other;  
Query Match 1.9%; Score 78.6; DB 19; Length 5285;  
Best Local Similarity 51.1%; Pred. No. 1.7e-08;  
Matches 408; Conservative 0; Mismatches 329; Indels 62; Gaps 7;

|      |                                                                 |    |
|------|-----------------------------------------------------------------|----|
| 1159 | TAATGGTGGCAGTTTGAAGATCCTCGTGAATGACAGACGATGCAATTTATGGACGTGATGTT  | QY |
| 1160 |                                                                 |    |
| 3225 | TATGTGATTTAGTTTGGAGATGCTCGTGAATGATGATGCAATTTATGCCCGTGTGATGTT    | Db |
| 1219 | TATGATTTTGATGGGTGTCGACTTCGGGTGTAGTAAA-----CGCATGATGAGAAAG       | QY |
| 3165 | TATGACTTTGATGGGCATCAITTTACGGGTGTTGTTAATCATCAAAATTTCAGAAATTTTACA | Db |
| 1268 | CTAGCTTAATTTCTGTAAATTTCTTGTAAGAGTGTTATCTTTTGTC--TGAATGTTTTTAGTT | QY |
| 3105 | CAAACTTTGTATTGTTGACTGCTGATGACCAACATAATTTATGTGTGTTGGCCATTGCAGGT  | Db |
| 1327 | TGAGATTGCACATGCTGGTCTGATGATTTTCAC---CATCAGTTGATAGTATGACAGACGAG  | QY |
| 3045 | GGAACTAGCTCATGTGTGGAGCGGTTTCATCATGATGACACGCGTAGTTATATAGTGGTCG   | Db |
| 1384 | CTACAGTGGAGCCGTGC-----ACCTCAAGACGCTCTGAC                        | QY |
| 2985 | TGCTCGTGGCGGTCGTGTGGTGGTGACGGTGGTGGTCTGTAACGTGGACCATCTAGGAG     | Db |
| 1421 | TACCGCGGTTGTAGAGTCTTCTCGATGTGTATTGTTGGTGTGTGTAATAATTTTATAATT    | QY |
| 2925 | ATCAGAGTACCGCGGTACATATGATATGTTTAGCTTTTAGTGTATTCATAGTTTTAGGAG    | Db |
| 1481 | TGAAAACCTCAITTTTACTACCTAAACATGCTAGTGGTTGTGACCGGATTAACCGCCTTCTGC | QY |
| 2865 | AAATCACTAAATTAGATTTTCCACTATACAGTTGTAGTGTGAGGTTTGCCTTCATCTGC     | Db |
| 1541 | TTGCTGGCAGGACCTTTAAGTTAAGGACACATATAGTCTTTTCTGAAATGTTGGTTTC      | QY |
| 2805 | GTCTTGGCAAGACCTCAAGTGATTAACACAGTTGTGGACAGATTC-----CTCAITTTGTCG  | Db |
| 1601 | TCTATATCATGTTTGGATTTTATCTTTTCTGAAATGATGTTATTTGCTATTATTCGGGT     | QY |
| 2750 | TTCCCTATATTTTCTACGATTTCTGAATGTGACATATCACTTTGTTT-----            | Db |
| 1661 | GATTAGGATCACATGCGAAAGCTGGAGATGCTGCTTCTCTGAACTTTTCCCTGACCGT      | QY |
| 2703 | -TGAAGGATCATGCTGTTAAAGGAGAGAGTTGTTGTTTCTCAAGTGTTCTGATGTT        | Db |
| 1721 | AAAGGTGAGTTGACATTCGATAGTTTGGATAAGCTTTTGTATGTGTTAGTAAATTA        | QY |
| 2644 | AGAGGTAAATCCTCATGCTTTACCCCAAGAAATAGAGATGTGCTAAAGCTGTTTGGGTTTT   | Db |
| 1781 | GTCTTTTGGAGGAGAAATAGTGTTAAGCATCTGAACTGCTAAA-----CTCACATTCAG     | QY |
| 2584 | CTGCGCTTTAAATTCCTGTTTTTAACCTAGTCTGTCATGCTTTAAAAAATCATATCGGGAC   | Db |







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 17:00:12 ; Search time 7850 Seconds

(without alignments)  
12520.675 Million cell updates/sec

Title: US-10-014-927-18

Perfect score: 4044

Sequence: 1 agacaaagatcttacttct.....atgattatcttaactaac 4044

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST:

1: em\_estba:\*\*

2: em\_estum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_mam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rod:\*\*

26: em\_gss\_phg:\*\*

27: em\_gss\_vrl:\*\*

28: gb\_ges1:\*\*

29: gb\_ges2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 571.4 | 14.1        | 1020   | 28    | Bi0994 F7G19-Sp6.3 |
| C 2        | 529   | 13.1        | 867    | 28    | B08242 F23J11-Sp6  |
| C 3        | 320.4 | 7.9         | 928    | 28    | Bi0052 F23J11-Sp6  |
| 4          | 309   | 7.6         | 973    | 28    | Bi1569 F3E19-Sp6 I |

|      |       |     |      |    |          |
|------|-------|-----|------|----|----------|
| C 5  | 305   | 7.5 | 445  | 9  | AV794356 |
| C 6  | 272.4 | 6.7 | 510  | 9  | AV531252 |
| C 7  | 265   | 6.6 | 276  | 29 | BZ358568 |
| 8    | 263.4 | 6.5 | 1044 | 28 | Bi11684  |
| 9    | 263   | 6.5 | 925  | 28 | Bi12280  |
| 10   | 260.2 | 6.4 | 340  | 29 | BZ592759 |
| 11   | 247.8 | 6.1 | 699  | 28 | Bi12780  |
| C 12 | 246.4 | 6.1 | 791  | 29 | BZ471603 |
| C 13 | 243.4 | 6.0 | 298  | 29 | AL950061 |
| C 14 | 235.6 | 5.8 | 300  | 9  | AV532256 |
| C 15 | 227.4 | 5.6 | 439  | 9  | AI998038 |
| C 16 | 223.8 | 5.5 | 694  | 28 | BZ065015 |
| C 17 | 209.8 | 5.2 | 691  | 28 | BZ065073 |
| C 18 | 207.2 | 5.1 | 467  | 9  | AW004187 |
| C 19 | 207.2 | 5.1 | 676  | 29 | BZ433601 |
| C 20 | 206.6 | 5.1 | 293  | 29 | AL950047 |
| C 21 | 203.4 | 5.0 | 277  | 14 | T42588   |
| C 22 | 202.4 | 5.0 | 411  | 14 | R65514   |
| C 23 | 188   | 4.6 | 199  | 29 | BZ381462 |
| C 24 | 187   | 4.6 | 762  | 28 | BZ062593 |
| C 25 | 178.8 | 4.4 | 741  | 28 | BH589916 |
| C 26 | 165.4 | 4.1 | 186  | 29 | BZ381461 |
| C 27 | 165.4 | 4.1 | 439  | 9  | AI994618 |
| C 28 | 132.8 | 3.3 | 311  | 14 | Z18005   |
| C 29 | 117   | 2.9 | 604  | 9  | AV826310 |
| C 30 | 113.8 | 2.8 | 675  | 9  | AV821889 |
| C 31 | 109.2 | 2.7 | 149  | 29 | AL757304 |
| C 32 | 98.6  | 2.4 | 361  | 28 | BH757020 |
| C 33 | 97.8  | 2.4 | 571  | 28 | AQ962088 |
| C 34 | 96.4  | 2.4 | 567  | 28 | AQ962089 |
| C 35 | 95.2  | 2.4 | 758  | 28 | BH540746 |
| C 36 | 90.8  | 2.2 | 762  | 28 | BH559250 |
| C 37 | 90.4  | 2.2 | 826  | 29 | BZ471593 |
| C 38 | 89.2  | 2.2 | 126  | 29 | AL757303 |
| C 39 | 86.6  | 2.1 | 724  | 28 | BH479317 |
| C 40 | 81.4  | 2.0 | 1078 | 29 | CNS014EL |
| C 41 | 80.6  | 2.0 | 997  | 14 | CD247514 |
| C 42 | 79.8  | 2.0 | 741  | 29 | CNS007YN |
| C 43 | 78    | 1.9 | 1132 | 29 | CC248307 |
| C 44 | 77.6  | 1.9 | 975  | 13 | BU960792 |
| C 45 | 76.6  | 1.9 | 1043 | 14 | CD385012 |

## ALIGNMENTS

RESULT 1  
Bi0994  
LOCUS F7G19-Sp6.3 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence.  
DEFINITION F7G19-Sp6.3 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence.  
ACCESSION Bi0994  
VERSION Bi0994.1 GI:2092116  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1 (bases 1 to 1020)  
AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.  
TITLE BAC End Sequences at ATGC  
JOURNAL Unpublished  
COMMENT F7G19-Sp6.2, F7G19-T7, F7G19-Sp6, F7G19-T7.3, F7G19-T7.4, F7G19-T7.3, F7G19-Sp6.1  
Arabidopsis Thaliana Genome Center  
University of Pennsylvania  
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104  
Tel: 215-898-9384  
Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu  
Seq primer: Sp6  
Class: BAC ends  
High quality sequence start: 81  
High quality sequence stop: 969.  
Location/Qualifiers

FEATURES

Source  
1..1020  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="F7G19"  
/sex="hermaphrodite"  
/clone\_lib="IGF"  
/note="Vector: BelobACII; Site\_1: EcoRI; Site\_2: EcoRI;  
Produced by Thomas Altmann"

BASE COUNT 283 a 234 c 166 g 331 t 6 others

Query Match 14.1%; Score 571.4; DB 28; Length 1020;  
Best Local Similarity 99.0%; Pred. No. 2.4e-73;  
Matches 607; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

QY 3433 TCGAATTCCTCAGTTTCACCTGTGGTAGTCTAAAGCTGAACCTCTCTTAATTCACAAT 3492  
Db 69 TCGAATTCCTCAGTTTCACCT---GTAAGTCTAAAGCTGAACCTCTCTTAATTCACAAT 125  
QY 3493 CCAT-GTGTGTTTAAATACCTGCTCACTTGGTGTCTTCAATCAACACCAACTAA 3551  
Db 126 CCATGGTGTGTTTAAATACCTGCTCACTTGGTGTCTTCAATCAACACCAACTAA 185  
QY 3552 CGAATCATGAGACAGACTATAAATTTGAAGAGTCTGTAGAAGCAGTAGTCTCCACAA 3611  
Db 186 CGAATCATGAGACAGACTATAAATTTGAAGAGTCTGTAGAAGCAGTAGTCTCCACAA 244  
QY 3612 CCTGTGTGCACTAAATAATCGCTCTCCAAAGTGTTCAGCAACATAATCTACCTGTGTC 3671  
Db 245 CCTGTGTGCACTAAATAATCGCTCTCCAAAGTGTTCAGCAACATAATCTACCTGTGTC 304  
QY 3672 ATGTGTTATCATTTCTTCTCTTAACGGTATTACATATTATGTTTGCAGGTGATATC 3731  
Db 305 ATGTGTTATCATTTCTTCTCTTAACGGTATTACATATTATGTTTGCAGGTGATATC 364  
QY 3732 TGGTGAATAATGAAACCTGGCACTGGCTGACCGAATGCTCTCAAGCTTCTCAGGCTC 3791  
Db 365 TGGTGAATAATGAAACCTGGCACTGGCTGACCGAATGCTCTCAAGCTTCTCAGGCTC 424  
QY 3792 CACTGCTAATAGAAATTTGATTCGGATTGGGATTATTATAGTGTCTTTTGTATGGGAC 3851  
Db 425 CACTGCTAATAGAAATTTGATTCGGATTGGGATTATTATAGTGTCTTTTGTATGGGAC 484  
QY 3852 GACCAATATGCTTTCTAGTTTATGTTGTAACCTGGAATGGTCTGTTATTTGTGCAAT 3911  
Db 485 GACCAATATGCTTTCTAGTTTATGTTGTAACCTGGAATGGTCTGTTATTTGTGCAAT 544  
QY 3912 AAAAGCGGAAACTCTGCTCGCTGCATATAAAGTTCATCAGCAATTTGTTGGTG 3971  
Db 545 AAAAGCGGAAACTCTGCTCGCTGCATATAAAGTTCATCAGCAATTTGTTGGTG 604  
QY 3972 TGGTGAGTGTTCATACATATACATTTACATTAACACTACTGGTGTCTTTTATGATTA 4031  
Db 605 TGGTGAGTGTTCATACATATACATTTACATTAACACTACTGGTGTCTTTTATGATTA 664  
QY 4032 TCTTAAACTAAAC 4044  
Db 665 TCTTAAACTAAAC 677

RESULT 2  
B08242/c  
LOCUS F23J11-Sp6.1 IGF Arabidopsis thaliana genomic clone F23J11, genomic  
DEFINITION survey sequence.  
867 bp DNA linear GSS 14-MAY-1997

ACCESSION

VERSION B08242.1 GI:2089364  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana  
Arabidopsis thaliana  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and  
Ecker, J.  
BAC End Sequences at ATGC

TITLE

Unpublished

JOURNAL

COMMENT Other GSSs: F23J11-Sp6, F23J11-T7.1, F23J11-T7

CONTACT

Contact: Ecker, J.  
Arabidopsis Thaliana Genome Center  
University of Pennsylvania  
Dept. of Biology, University of Pennsylvania, Philadelphia, PA  
19104

Tel: 215-898-9384  
Fax: 215-898-9780  
Email: jecker@atgenome.bio.upenn.edu

Seq primer: Sp6  
Class: BAC ends  
High quality sequence start: 48  
High quality sequence stop: 549.

FEATURES

source

1..867  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="F23J11"  
/sex="hermaphrodite"  
/clone\_lib="IGF"  
/note="Vector: BelobACII; Site\_1: EcoRI; Site\_2: EcoRI;  
Produced by Thomas Altmann"  
BASE COUNT 290 a 207 c 153 g 209 t 8 others  
ORIGIN

Query Match 13.1%; Score 529; DB 28; Length 867;  
Best Local Similarity 82.8%; Pred. No. 3.6e-67;  
Matches 684; Conservative 0; Mismatches 132; Indels 10; Gaps 7;

QY 2647 TTCTTCCTTTGTCATGTTTCTATGTCATGCAACCTCCATATGGGTAAAGGTACT 2706  
Db 838 TTCTTCCTTTGTCATGTTTCTATGTCATGCAACCTCCATATGGGTAAAGGTACT 779  
QY 2707 CCTTGTGTTGGGATTACCAAGTTCCTTTTCTTACGTGAATGTTGTTTGT 2766  
Db 778 CCTTGTGTTGGGATTACCAAGTTCCTTTTCTTACGTGAATGTTGTTTGT 719  
QY 2767 ATGTTTTCAGTTCTTTGACAGAGATGCTCCCATCATATTAGTCCCTTTCTCTTTG 2826  
Db 718 AGTTTTCAGTTCTTTGACAGAGATGCTCCCATCATATTAGTCCCTTTCTCTTTG 659  
QY 2827 TGTGTTCTTCTGAGTGTTCCTTCTGATAAAGCTTACTTCTTAATCTTTTCCAGC 2886  
Db 658 GGTTCATCTCTTNGGAGGTTCCTTCTGATGAAGTAAACTTCTAAATTTTTCCTCAGC 599  
QY 2887 GACGTGTAATTTATACGTATCAACCTCAATCCGACTATATTAAAGACACTTAG 2946  
Db 598 GACGTGTAATTTATACGTATCAACCTCAATCCGACTATATTAAAGACACTTAG 540  
QY 2947 CTAGATGTT---CACTTTTGAATAATTTATTTCTATCGGAAGGGGCA--GATGATTTCTG 3001  
Db 539 GTAGATGTTTCACTTTTGAATAATTTATTTCTATCGGAAGGGGAGATGATTTCTG 480  
QY 3002 AGGCATGTCCTCGATCATTTACCGTGATTAACAATTTGATTCGGTTGTTG-TTCTAGAT 3060  
Db 479 AGGCATGTCCTCGATCATTTACCGTGATTAACAATTTGATTCGGTTGTTG-TTCTAGAT 420  
QY 3061 CTGGCTCACTGCTAGCAGCTGGGATGGAATGGAATCTAGATGGGTCACTAGATGCTTGG 3120

[illegible]

|            |                                                                                    |        |            |
|------------|------------------------------------------------------------------------------------|--------|------------|
| RESULT 3   |                                                                                    |        |            |
| B10052/c   |                                                                                    |        |            |
| LOCUS      | B10052                                                                             | 928 bp | DNA linear |
| DEFINITION | F23J11-Sp6 IGF Arabidopsis thaliana genomic clone F23J11, genomic survey sequence. |        |            |

|           |                                                                     |
|-----------|---------------------------------------------------------------------|
| ACCESSION | B10052                                                              |
| VERSION   | B10052.1                                                            |
| KEYWORDS  | GI:2091194                                                          |
| SOURCE    | GSS.                                                                |
| ORGANISM  | Arabidopsis thaliana (thale cress)                                  |
|           | Arabidopsis thaliana                                                |
|           | Arabidopsis thaliana                                                |
|           | Eukaryota; Viridiplanetae; Streptophyta; Embryophyta; Tracheophyta; |
|           | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids |
|           | ; eucotids II; Brassicales; Brassicaceae; Arabidopsids.             |
| REFERENCE | 1 (bases 1 to 928)                                                  |
| AUTHORS   | Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and  |
|           | Ecker,J.                                                            |

| TITLE   | BAC End Sequences at ATGC                                      |
|---------|----------------------------------------------------------------|
| JOURNAL | Unpublished                                                    |
| COMMENT | Other GSS: F23J11-SP6.1, F23J11-T7.1, F23J11-T7                |
|         | Contact: Ecker J.                                              |
|         | Arabidopsis Thaliana Genome Center                             |
|         | University of Pennsylvania                                     |
|         | Dept. of Biology, University of Pennsylvania, Philadelphia, PA |
|         | 19104                                                          |

tel: 215-898-2384  
fax: 215-898-8780  
Email: [jecker@genome.bio.upenn.edu](mailto:jecker@genome.bio.upenn.edu)  
Seq primer: Sp6  
Class: BAC ends  
High quality sequence start: 105  
High quality sequence stop: 133.

```

FEATURES
  source
    high quality sequence: 100%
    location/Qualifiers
      1. 928
        /organism="Arabidopsis thaliana"
        /mol_type="genomic DNA"
        /strain="Columbia"
        /db_xref="taxon:3702"
        /clone="F03J11"
        /sex="hermaphrodite"
        /clone_lib="IGF"
        /note="vector: BEL08AC11; Site_1: EcoRI; Site_2: EcoRI;
        Produced by Thomas Altmann"
        268 a 249 c 201 g 199 t 11 others
BASE COUNT

```

|            |                             |       |       |       |           |
|------------|-----------------------------|-------|-------|-------|-----------|
| BASE COUNT | 268 a                       | 249 c | 201 g | 199 t | 11 others |
| ORIGIN     | Produced by Thomas Arltmann |       |       |       |           |

| Query Match                                                      | 7.9%; Score 320.4; DB 28; Length 928;                                   |
|------------------------------------------------------------------|-------------------------------------------------------------------------|
| Best local similarity                                            | 76.2%; Pred. No. 6e-37;                                                 |
| Matches 552; Conservative 0; Mismatches 151; Indels 21; Gaps 12; |                                                                         |
| QY                                                               | 2739 TTCTTACACGTGAAGTGTGTTTGTTTTATGTTTTGAGTCTCTTGACAGAGATGCTCCCAT 2798  |
| DB                                                               | 797 TTTTCACTTCAAAAGGAGTGGTTGTTTTTAATGTCAGGGTTCGGACGAGGCTCCCAT 738       |
| QY                                                               | 2799 CATATTAGTCCTTTTCCTTTTCTTTGTCTGCTTCTTCTTCGGATGTTTCCCTTCGTATA 2858   |
| DB                                                               | 737 CAAATT---ATCCTTTCCCTTCTCCTCGGGTCG---TGTAAATGGGGAAATCCGTTCCGTA 684   |
| QY                                                               | 2859 AAGCTTACTCTTAACTTTTTCCTTTTTCAGACGCGTGAATTTATTCGTATCAACCTCAATA 2918 |
| DB                                                               | 683 AAGCTTGATTC-TCACTTTTTCAGCGACGTGA--TAAATAGCGTATCATGCTCAGGA 627       |
| QY                                                               | 2919 TCGACCTATAATTTAAAGAACACTTAGCTAGATGTTCACTTTTGAAAAATTTATTCTAT 2978   |
| DB                                                               | 626 TCGACCTATAA--TAAACGACACATAGCTAGAGTTTCACTCTTGAAAGTTTATGTCAT 569      |
| QY                                                               | 2979 CGGAAGGGGCGAGT-GATTTCTGAGGAGTCTCTCGATCATTTACCGTGATTTACAATTG 3037   |
| DB                                                               | 568 CGGAAGGGTAAAGATGGATTTCTGAGGCAATGCTCTCGATCATTAACGTAGTGACTATGG 509    |
| QY                                                               | 3038 TATTGGGTTGTTGT--TTCTAGATCTGGCTCACTGCTACGAGCTGGGATTTGGATCTAGA 3096  |
| DB                                                               | 508 TATCGGTTGCTTTGTTTCTAGATCAGGCTCACTGCTACGAGCTGGGATTTGGATCTAGA 449     |
| QY                                                               | 3097 TGGGTCATCTAGATGGATTTCTTGGACTGATTTACAAAGCTGGATTAGATGAAC-TGAA 3155   |
| DB                                                               | 448 TGGTTTCATCAAGATGGATTCGTGGACTGGATCCACAGAGCTGGCTTCGCTGAACTTGAA 389    |
| QY                                                               | 3156 CTTCTGTTT-TACGGTCTGGTCTGGTCTGGTACTCGGCGGTATCAGCTGTAGGATCTGA 3214   |
| DB                                                               | 388 CTTCTGCTCTACGGTCTGGTCCGATGTCGTACTCGGCGGTATCAGTTGTAGGATCTGA 329      |
| QY                                                               | 3215 TCGCAAAGTTTGGACTATGATTACTCTGATTCCTCAATATATTTATCTTTTTCACAAAT 3274   |
| DB                                                               | 328 TGGCAAAGCTCTGGACATTGATTACTCTGATTTCGTCGATATATTGATCTTTTTCACAA 269     |
| QY                                                               | 3275 GTGGATCTGTGTAGTTCTTTTCTAGACAGCATTTAAAGCTCCGGGACTAGATGGGA 3334      |
| DB                                                               | 268 GTGGACTCTGGGTCG-GTTATATATTAGACAGCATTTTAAAGCTCCGGGACTAGATGGCG 210    |
| QY                                                               | 3335 GATGGTCAGTAAATTTCTTTGTATTGCCACACTTACATGGGGTTTTTCGCTTCGTGGA 3394    |
| DB                                                               | 209 TGT-----CAGTCAATTTGTTTGTATGCCACATGATGGGGTTTTT-GGTCTTGTGGA 155       |
| QY                                                               | 3395 GGTCCCAATCAAGATCAAAATCAAGATCAAGATCAAGATCGAATTTCTCCAGTTTCACTTG 3454 |
| DB                                                               | 154 TGTCCCACTCGAGATCTAAGTCGAGATCAAGATCAAGATCGAATTCGAGCTCGTGAACCG 95     |
| QY                                                               | 3455 TGGT 3458                                                          |
| DB                                                               | 94 GGAT 91                                                              |

|            |                                                                                  |        |     |                 |
|------------|----------------------------------------------------------------------------------|--------|-----|-----------------|
| RESULT 4   |                                                                                  |        |     |                 |
| B11569     |                                                                                  |        |     |                 |
| LOCUS      | B11569                                                                           | 973 bp | DNA | linear          |
| DEFINITION | F3E19-Sp6 IGF Arabidopsis thaliana genomic clone F3E19, genomic survey sequence. |        |     | GSS 14-MAY-1997 |

B11569  
 ACCESSION  
 B11569.1 GI:2092690  
 VERSION  
 GSS.  
 KEYWORDS  
 SOURCE  
 Arabidopsis thaliana  
 Arabidopsis thaliana  
 ORGANISM

KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM

*Arabidopsis thaliana*  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 973)

REFERENCE  
Feng, J., Dewar, K., Buehler, B., Kim, C., Li, Y., Shinn, P., Sun, H. and  
Ecker, J.



QY 4015 GGTGCTCTTTATG 4027  
Db 13 GGTGCTTTATG 1

RESULT 6  
LOCUS AV531252/c 510 bp mRNA linear EST 01-SEP-2000  
DEFINITION AV531252 Arabidopsis thaliana flower buds Columbia Arabidopsis thaliana cDNA clone FB018C01F 3', mRNA sequence.

ACCESSION AV531252  
VERSION AV531252  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 510)  
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries  
JOURNAL DNA Res. 7, 175-180 (2000)  
MEDLINE 20363093  
PUBMED 10907847

COMMENT Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp; URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
source  
1..510  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="FB018C01F"  
/tissue\_type="flower buds"  
/clone\_lib="Arabidopsis thaliana flower buds Columbia"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 175 a 110 c 99 g 126 t

ORIGIN  
Query Match 6.7%; Score 272.4; DB 9; Length 510;  
Best Local Similarity 99.3%; Pred. No. 6.9e-30;  
Matches 284; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3174 GTCTGGTCTGTCTCCGCGGTATCAGCTGTAGGATCTGATCGCAAGTTTGGACTAT 3233  
Db 510 GTCTGGTCTGTCTCCGCGGTATCAGCTGTAGGATCTGATCGCAAGTTTGGACTAT 451

QY 3234 GATTACTGTATCTCTCAATATTATCTTTTTCACCAATAGTGGANTCTGTGTGAGTT 3293  
Db 450 GATTACTGTATCTCTCAATATTATCTTTTTCACCAATAGTGGANTCTGTGTGAGTT 391

QY 3294 CTTTCTAGGACAGATTATAGCTCCGGGACTAGATCGGAGATGGTCAGTAATTTCTT 3353  
Db 390 CTTTCTAGGACAGATTATAGCTCCGGGACTAGATCGGAGATGGTCAGTAATTTCTT 331

QY 3354 TCTTATGCCACACTTACATGGGTTTTCGGTCTTCTGCTGCGAGTCCCAATCAAGATCAA 3413  
Db 330 TCTTATGCCACACTTACATGGGTTTTCGGTCTTCTGCTGCGAGTCCCAATCAAGATCAA 272

QY 3414 ATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 3459  
Db 271 ATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 226

RESULT 7  
LOCUS BZ358568  
DEFINITION SALK\_132986.49.85.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK\_132986.49.85.x, genomic survey sequence.

ACCESSION BZ358568  
VERSION BZ358568.1  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 276)  
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,I., Shinn,P., Zimmermann,J. and Ecker,J.R.  
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome  
JOURNAL Unpublished  
COMMENT Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGnAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu

FEATURES  
source  
1..276  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="SALK\_132986.49.85.x"  
/clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
/notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

BASE COUNT 70 a 56 c 48 g 102 t

ORIGIN  
Query Match 6.8%; Score 265; DB 29; Length 276;  
Best Local Similarity 99.6%; Pred. No. 1e-28;  
Matches 276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 658 CCAACGAAAGGTTGTCTCTCTTCCAGAACATCTAACGTTTCTCGACATCTTCT 717  
Db 1 CCAACGAAAGGTTGTCTCTCTTCCAGAACATCTAACGTTTCTCGACATCTTCT 60

QY 718 TCTTCTCTTCTCGAAATTTATTTTCCAGTAATCAATTTCTTCTCTTAGATTTTAC 777  
Db 61 TCTTCTCTTCTCGAAATTTATTTTCCAGTAATCAATTTCTTCTCTTAGATTTTAC 120

QY 778 AGGAACATAATTTTCTCTCTGAGGATCAGATGAGTAGCGGATGGAATCGTACGATCTAC 837  
Db 121 AGGAACATAATTTTCTCTCTGAGGATCAGATGAGTAGCGGATGGAATCGTACGATCTAC 180

QY 838 GTTGGGAATTTGCTCGAGATATTCGAAAGTGTGAGGTTGAAGATCTCTTCTACAAGTT 897  
Db 181 GTTGGGAATTTGCTCGAGATATTCGAAAGTGTGAGGTTGAAGATCTCTTCTACAAGTT 240

QY 898 TGAAGATTTCTCTCTTCTCGATATAAATTAATTGAATT 934  
Db 241 TGAAGATTTCTCTCTTCTCGATATAAATTAATTGAATT 276

RESULT 8  
LOCUS B11684

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                         |                  |           |              |                 |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|------------------|-----------|--------------|-----------------|
| LOCUS      | B11684                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                         | 1044 bp          | DNA       | linear       | GSS 14-MAY-1997 |
| DEFINITION | F7G19-Sp6.2 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence.                                                                                                                                                                                                                                                                                                                                                                                         |                                                         |                  |           |              |                 |
| ACCESSION  | B11684                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                         |                  |           |              |                 |
| VERSION    | B11684.1                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | GI:2092807                                              |                  |           |              |                 |
| KEYWORDS   | GSS.                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                         |                  |           |              |                 |
| SOURCE     | Arabidopsis thaliana (thale cress)                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                         |                  |           |              |                 |
| ORGANISM   | Arabidopsis thaliana                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                         |                  |           |              |                 |
|            | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsia.                                                                                                                                                                                                                                                                              |                                                         |                  |           |              |                 |
| REFERENCE  | 1 (bases 1 to 1044)                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                         |                  |           |              |                 |
| AUTHORS    | Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and Ecker,J.                                                                                                                                                                                                                                                                                                                                                                                                |                                                         |                  |           |              |                 |
| TITLE      | BAC End Sequences at ATGC                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                         |                  |           |              |                 |
| JOURNAL    | Unpublished                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                         |                  |           |              |                 |
| COMMENT    | Other GSSs: F7G19-T7.2, F7G19-T7.1, F7G19-T7.4, F7G19-Sp6.3, F7G19-T7.3, F7G19-T7, F7G19-Sp6, F7G19-Sp6.1<br>Contact: Ecker J.<br>Arabidopsis Thaliana Genome Center<br>University of Pennsylvania<br>Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104<br>Tel: 215-898-9384<br>Fax: 215-898-8780<br>Email: jecker@atgenome.bio.upenn.edu<br>Seg primer: Sp6<br>Class: BAC ends<br>High quality sequence start: 207<br>High quality sequence stop: 432. |                                                         |                  |           |              |                 |
| FEATURES   | source                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Location/Qualifiers                                     |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1..1044                                                 |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /organism="Arabidopsis thaliana"                        |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /mol_type="genomic DNA"                                 |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /strain="Columbia"                                      |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /db_xref="taxon:3702"                                   |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /clone="F7G19"                                          |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /sex="hermaphrodite"                                    |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /clone_lib="IGF"                                        |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | /note="Vector: BelorACII; Site_1: EcoRI; Site_2: EcoRI; |                  |           |              |                 |
|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Produced by Thomas Altman"                              |                  |           |              |                 |
| BASE COUNT | 297 a                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 228 c                                                   | 107 g            | 341 t     | 71 others    |                 |
| ORIGIN     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                         |                  |           |              |                 |
|            | Query Match                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 6.5%;                                                   | Score 263.4;     | DB 28;    | Length 1044; |                 |
|            | Best Local Similarity                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 71.0%;                                                  | Pred. No. 1e-28; |           |              |                 |
|            | Matches 370;                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Conservative 0;                                         | Mismatches 145;  | Indels 6; | Gaps 4;      |                 |
| Qy         | 3409 TC AAAATCAAGATCAAGATCAAGATCGAATTCCAGTTTCACCTGGTGGTAAGCTAAAA                                                                                                                                                                                                                                                                                                                                                                                                           | 3468                                                    |                  |           |              |                 |
| Db         | 111 TC ATATATGATCCNGGTACCAAACCTCAAATCTCCAATTTTCACCT--GTAACCTCAAAA                                                                                                                                                                                                                                                                                                                                                                                                          | 167                                                     |                  |           |              |                 |
| Qy         | 3469 GCTGAACCTTCTTTAATTCACAATCCAT-GTGTTTGTTTAAATACCTGCTCATCTTGGTT                                                                                                                                                                                                                                                                                                                                                                                                          | 3527                                                    |                  |           |              |                 |
| Db         | 168 NCTAAACCTTCTTTAAATTCACAATCCCATNGTGTGTGTTGTTTAAATACCTGCTCATCTTGGTT                                                                                                                                                                                                                                                                                                                                                                                                      | 227                                                     |                  |           |              |                 |
| Qy         | 3528 GTTCTTCAATCAACACCACCTTAACGAAATCATGAGACAGACTATAAAATTTGAAGATC                                                                                                                                                                                                                                                                                                                                                                                                           | 3587                                                    |                  |           |              |                 |
| Db         | 228 GTTCTTCAATCAACACCACCTTAACGAAATCATGAGACAGACTATAAAATTTGAAGATC                                                                                                                                                                                                                                                                                                                                                                                                            | 287                                                     |                  |           |              |                 |
| Qy         | 3588 TGTAGAACGACTAGGTCTCACCAACCTCTGTGTGCATATAAAATTCGCCCTCTCCAAGTGT                                                                                                                                                                                                                                                                                                                                                                                                         | 3647                                                    |                  |           |              |                 |
| Db         | 288 T-FAGAACGACTANGTCTACCAACCTCTGTGTGCATATAAAATTCGCCCTCTCCAAGTGT                                                                                                                                                                                                                                                                                                                                                                                                           | 346                                                     |                  |           |              |                 |
| Qy         | 3648 TCAGCACATAATCTACTCTGTGCATGTGTATCATTTCTTCTTTAAACGGTATTAC                                                                                                                                                                                                                                                                                                                                                                                                               | 3707                                                    |                  |           |              |                 |
| Db         | 347 TCAGCACATAATCTACTCTGTGCATGTGTATCATTTCTTCTTCTTAAACGGTATTAC                                                                                                                                                                                                                                                                                                                                                                                                              | 406                                                     |                  |           |              |                 |
| Qy         | 3708 ATATTATGTTTTGCAGGTGATATCTGGTTGAAATGAAGACTGCCACTGGCTGTACCCG                                                                                                                                                                                                                                                                                                                                                                                                            | 3767                                                    |                  |           |              |                 |
| Db         | 407 ATATTATGTTTTGCAGGTGATATCAANNNNNANNTNANNCNNNNNNNNNNNNNNNNNNC                                                                                                                                                                                                                                                                                                                                                                                                            | 466                                                     |                  |           |              |                 |
| Qy         | 3768 AATCGTCTCAAGCTTCTFCAGGCTCCACTGCTGAATAGAAATTTGATTCGATTTTCGGGATTAT                                                                                                                                                                                                                                                                                                                                                                                                      | 3827                                                    |                  |           |              |                 |

|    |      |                                                                |      |
|----|------|----------------------------------------------------------------|------|
| Qy | 3605 | TCACCAACTCTGTGTGCGACTAAAAAATCGCCTCTCCAGTGGTTTCAGACAATAAATCTAC  | 3664 |
| Db | 281  | TCACCAACTCTGTGTGCGACTAAAAATCGCCTCTCCAGTGGTTTCAGACAATAAATCTAC   | 340  |
| Qy | 3665 | CTCTGCATGTGGTATTCAATTCCTTCTCTTAACGGTATTACATATTATGTTTTTGACGG    | 3724 |
| Db | 341  | ATCTGTGATGTGTTATCATNTCTTCTCTTAACGGTATTACATATTATGTTTGACGG       | 400  |
| Qy | 3725 | TGATATCTGTTTGAANAATGAAAACCTGGCACCTGGCTGTACCGGAATCGTCTCAAGCTTCT | 3784 |
| Db | 401  | TGATAAGAGNNNNNANNNNNNNNNNNNNGNTNTNCCANNCNNCCNNCCNNCCNC-        | 459  |
| Qy | 3785 | CAGCTCCACTGCTAAATAGAATTTGATTCCGATTTGGGATTATTATACTGGTCTCTTTGT   | 3844 |
| Db | 460  | --NGGCCCNCCGNNNNCACAAATGNACTTCGATCTGGGAATA-TCTCCGGTCTCTTTT     | 516  |
| Qy | 3845 | ATGGGACGACCA                                                   | 3856 |
| Db | 517  | TGGGAACAACCA                                                   | 528  |

RESULT 10  
BZ592759  
LOCUS BZ592759 340 bp DNA linear GSS 07-JAN-2003  
DEFINITION SALK\_029105.23.05.x Arabidopsis thaliana TDNA insertion lines.  
Arabidopsis thaliana genomic clone SALK\_029105.23.05.x, genomic  
survey sequence.

ACCESSION BZ592759  
VERSION BZ592759.1  
KEYWORDS GI:2753278  
GSS.

| SOURCE                             | ORGANISM                                                                                                                               |
|------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Arabidopsis thaliana (thale cress) | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids |
| Arabidopsis thaliana               | ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.                                                                                 |

REFERENCE  
1 (bases 1 to 340)  
AUTHORS  
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab  
, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P.,  
Zimmerman, J. and Ecker, J.R.  
TITLE  
A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome  
Unpublished  
Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGnAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of  
TDNA.

```

FEATURES
source
      Class: TDNA tagged.
      Location/Qualifiers
        1..340
          /organism="Arabidopsis thaliana"
          /mol_type="genomic DNA"
          /strain="Columbia 0"
          /db_xref="taxon:3702"
          /clone="SALK_029105.23.05.x"
          /clone_lib="Arabidopsis thaliana TDNA insertion lines"
          /notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"
          96 a 69 C 73 g 102 t
BASE COUNT

```

```

BASE COUNT      96 a      69 c      73 g      102 t
ORIGIN
Query Match      6.4%; Score 260.2; DB 29; Length 340;
Best Local Similarity 87.1%; Pred. No. 4.7e-28;
Matches 297; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

```

|    |     |                                                                 |     |
|----|-----|-----------------------------------------------------------------|-----|
| Qy | 449 | GATTGAAATTTGGTTCCGAATTTGTAATAACATAACAATATTCGCGTTTGAATGATAAG     | 508 |
| Db | 1   | GATTTCGAATTTTGGTTGGAATTTGTAATAACTAACGTACAATATTCGGGGTTGAATGATGAT | 60  |
| Qy | 509 | AAAAAACACATTCGATCCGGTTAGACAATATATTAAACAGGCCCATATAAACATATTCGGC   | 568 |
| Db | 61  | AAAAAACACATTCGATCCGGTTAGAACCATATATTAAACGGGCCCTTTAAAA-AGATCGGC   | 119 |
| Qy | 569 | CGATCTTTGATCAACTCGGCTATTTCATCGTTGTATACATCGCGCGCACAGGATTAATAATCC | 628 |
| Db | 120 | CGATCTTTGATCAACTCGGCTATTTCACGGTTGTATACATCGCGGGCGACAGATGAAGGCC   | 179 |
| Qy | 629 | AGTTCGGTTTATAAAGGATACTAGTTTCCAAACGAACGGTGGTTGTCTCTCTTCCAGA      | 688 |
| Db | 180 | GGTGGCGTTTATAAAGGATACTACTTTTCCAAACGAAACGAGGTTGTCTCTCTGTCAGA     | 239 |
| Qy | 689 | ACAATCTAACCGTTTCTCGAACATCTTCTTCTTCTTCTTCGAAATTAATTTTCCAGT       | 748 |
| Db | 240 | ACTATCTGAAGGCTTCTGACCACTCTCTTCTTCTCACTCTCTAGAAATATGCTTTCAGGG    | 299 |
| Qy | 749 | AATCAATTTCTTCTCTTAGATTTTTACAGGAACTAATTT                         | 789 |
| Db | 300 | AAACACTTTCTGCTCTTCTTAGATCTTTTAGGGAACTAAATTT                     | 340 |

|            |                                                                                    |        |     |        |
|------------|------------------------------------------------------------------------------------|--------|-----|--------|
| RESULT 11  |                                                                                    |        |     |        |
| B12780     |                                                                                    |        |     |        |
| LOCUS      | B12780                                                                             | 699 bp | DNA | linear |
| DEFINITION | F7G19-Sp6.1 IGF Arabidopsis thaliana genomic clone F7G19, genomic survey sequence. |        |     |        |

survey sequence.  
 B12780  
 VERSION B12780.1 GI:2093901  
 KEYWORDS GSS.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eumatosids II; Brassicales; Brassicaceae; Arabidopsiis.  
 1 (bases 1 to 699)  
 REFERENCE Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and  
 Eckert, J.

TITLE BAC End Sequences at ATGC  
 UNPUBLISHED  
 OTHER GSSs: F7G19-T7.2, F7G19-T7.1, F7G19-T7.4, F7G19-Sp6.3,  
 F7G19-T7.3, F7G19-Sp6.2, F7G19-T7, F7G19-Sp6  
 CONTACT: Ecker J.  
 Arabidopsis Thaliana Genome Center  
 University of Pennsylvania  
 Dept. of Biology, University of Pennsylvania, Philadelphia, PA  
 19104  
 Tel: 215-898-9384

```

FEATURES
    source
        high quality sequence BSCP: 474.
        Location/Qualifiers
            1..699
                /organism="Arabidopsis thaliana"
                /mol_type="genomic DNA"
                /strain="Columbia"
                /db_xref="taxon:3702"
                /clone="F7G19"
                /sex="hermaphrodite"
                /clone_lib="IGF"
                /note="Vector: BeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                Produced by Thomas Altmann"
            103 a 235 c 77 g 274 t 10 others
BASE COUNT

```

| BASE COUNT  | 103 A                                 | 235 C | 77 G | 274 T | 10 OTHERS |
|-------------|---------------------------------------|-------|------|-------|-----------|
| ORIGIN      |                                       |       |      |       |           |
| Query Match |                                       |       |      |       |           |
|             | 6.1%; Score 247.8; DB 28; Length 699; |       |      |       |           |



|                                                                              |                                                                     |                     |  |  |
|------------------------------------------------------------------------------|---------------------------------------------------------------------|---------------------|--|--|
| Best Local Similarity 94.5%; Pred. No. 2.2e-26;                              |                                                                     |                     |  |  |
| Matches 290; Conservative 0; Mismatches 12; Indels 5; Gaps 3;                |                                                                     |                     |  |  |
| QY                                                                           | 3426 ATCAAGATCGAATTCCTCCAGTTTCACCTGTGGTAAGTCTAAAGCTGAACCTCTTTTAAT   | 3485                |  |  |
| Db                                                                           | 90 ACCGAGCTCGAGTCTCTCCCTTTCACCTG---TAAATCTAAAGCTGAACCTCTTTTAAT      | 146                 |  |  |
| QY                                                                           | 3486 TCACAATCCAT-GTGTTTGTTTAAATACCTGCTCACCCTTTGGTTGTTCTTCAATCAACACC | 3544                |  |  |
| Db                                                                           | 147 TCACAATCCATGGTGTGTTTAAATACCTGCTCACCCTTTGGTTGTTCTTCAATCAACACC    | 206                 |  |  |
| QY                                                                           | 3545 AACCTTAACGAATCATGACAGACTATAAAATTTGAAGAGTCTGTAGACGACTAGGTC      | 3604                |  |  |
| Db                                                                           | 207 AACCTTAACGAATCATGACAGACTATAAAATTTGAATAGTCT-TAGAACGACTAGGTC      | 265                 |  |  |
| QY                                                                           | 3605 TCACCAACCTCTGTGTCACATAAAATCGCTCTCCAAAGTGTTCAGCAACATAATCTAC     | 3664                |  |  |
| Db                                                                           | 266 TCACCAACCTCTGTGTCACATAAAATCGCTCTCCAAAGTGTTCAGCAACATAATCTAC      | 325                 |  |  |
| QY                                                                           | 3665 CTCTGTGATGTTATCATTTCTTCTCTTAACGGTATATACATATATGTTTTCAGG         | 3724                |  |  |
| Db                                                                           | 326 CTCTGTGATGTTATCATTTCTTCTCTTAACGGTATATACATATATGTTTTCAGG          | 385                 |  |  |
| QY                                                                           | 3725 TGATATC 3731                                                   |                     |  |  |
| Db                                                                           | 386 TGATAAC 392                                                     |                     |  |  |
| RESULT 12                                                                    |                                                                     |                     |  |  |
| BZ471603/LOCUS                                                               |                                                                     |                     |  |  |
| BZ471603 791 bp DNA linear GSS 13-DEC-2002                                   |                                                                     |                     |  |  |
| BONJM74R BO 1.6.2_KB tot Brassica oleracea genomic clone BONJM74,            |                                                                     |                     |  |  |
| genomic survey sequence.                                                     |                                                                     |                     |  |  |
| ACCESSION BZ471603                                                           |                                                                     |                     |  |  |
| VERSION BZ471603.1 GI:26769878                                               |                                                                     |                     |  |  |
| KEYWORDS GSS.                                                                |                                                                     |                     |  |  |
| SOURCE Brassica oleracea                                                     |                                                                     |                     |  |  |
| ORGANISM Brassica oleracea                                                   |                                                                     |                     |  |  |
| REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |                                                                     |                     |  |  |
| Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids          |                                                                     |                     |  |  |
| ; eurosids II; Brassicales; Brassicaceae; Brassica.                          |                                                                     |                     |  |  |
| AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.         |                                                                     |                     |  |  |
| TITLE Whole genome shotgun sequencing of Brassica oleracea                   |                                                                     |                     |  |  |
| JOURNAL Unpublished                                                          |                                                                     |                     |  |  |
| COMMENT Other_GSSs: BONJM74TF                                                |                                                                     |                     |  |  |
| Contact: Chris Town                                                          |                                                                     |                     |  |  |
| TIGR                                                                         |                                                                     |                     |  |  |
| 9712 Medical Center Drive, Rockville, MD 20850, USA.                         |                                                                     |                     |  |  |
| Tel: 301-838-3523                                                            |                                                                     |                     |  |  |
| Fax: 301-838-0208                                                            |                                                                     |                     |  |  |
| Email: cdtown@tigr.org                                                       |                                                                     |                     |  |  |
| DNA is from a doubled haploid provided by Tom Osborn.                        |                                                                     |                     |  |  |
| Seq primer: TR                                                               |                                                                     |                     |  |  |
| Class: sheared ends.                                                         |                                                                     |                     |  |  |
| FEATURES                                                                     | source                                                              | Location/Qualifiers |  |  |
| 1..791                                                                       |                                                                     |                     |  |  |
| /organism="Brassica oleracea"                                                |                                                                     |                     |  |  |
| /mol_type="genomic DNA"                                                      |                                                                     |                     |  |  |
| /strain="TO1000DH3"                                                          |                                                                     |                     |  |  |
| /db_xref="taxon:3712"                                                        |                                                                     |                     |  |  |
| /clone="BONJM74"                                                             |                                                                     |                     |  |  |
| /clone_lib="BO 1.6.2 KB tot"                                                 |                                                                     |                     |  |  |
| /note="Vector: pBOS1; Site 1: BstXI; 1.6-2 kb sheared                        |                                                                     |                     |  |  |
| total DNA inserted into pBOS1 using BstXI linkers"                           |                                                                     |                     |  |  |
| BASE COUNT                                                                   | 295 a                                                               | 185 c 119 g 192 t   |  |  |
| ORIGIN                                                                       |                                                                     |                     |  |  |
| Query Match 6.1%; Score 246.4; DB 29; Length 791;                            |                                                                     |                     |  |  |
| Best Local Similarity 64.3%; Pred. No. 3.4e-26;                              |                                                                     |                     |  |  |
| Matches 491; Conservative 0; Mismatches 221; Indels 52; Gaps 6;              |                                                                     |                     |  |  |
| QY                                                                           | 1523 CGGATTACCGCCCTCTCTCTCGTGGCAGGACCTTAAGGTAAGGACACTATATAGTCTTT    | 1582                |  |  |

|                                                                             |      |                                                               |      |
|-----------------------------------------------------------------------------|------|---------------------------------------------------------------|------|
| Db                                                                          | 784  | CGGGATACCGCACTCTGCTTCTGCGCAAGACTTAAGGTAAGGATGTTACTTGTCTGTTT   | 725  |
| QY                                                                          | 1583 | TTCTCTGAATGTGTTGTTCTCTCT-----ATATCATGTTTGGATTATCTCTTTTCT      | 1632 |
| Db                                                                          | 724  | CATTTTCATGTTGTTTATTTGGATGTTGCGTCTCTAGTTCACTGTTCAATTTATTTGT    | 665  |
| QY                                                                          | 1633 | GAAATGATGTTATTTGCTTATTTACGGGTGATAGGATCACATGCGCAAGCTGGAGATGT   | 1692 |
| Db                                                                          | 664  | TCAGAGCTCATGTTTATTTGTTGGGGGATTAAGGATCACATGCGTAAAGCTGGAGATGT   | 605  |
| QY                                                                          | 1693 | CTGCTTCTCTAGGTTTTCCTGACCGTAAAGGTGAGT-----TGACATTCGATAGTTTGG   | 1748 |
| Db                                                                          | 604  | CTGCTTCTCTAGGTTTTCCTGACCGTAAAGGTGAGT-----TGACATTCGATAGTTTAA   | 545  |
| QY                                                                          | 1749 | ATACCTTTTTCATGATGTTTCTTGTAGTAAATAGTCTTTGTGAAGAGAGATAGGTGTTTAA | 1808 |
| Db                                                                          | 544  | ATAACCTTTCTGTAGATATATAAATAATTTTAGTGTCTGTGAAGAGATAGTCTAGTT     | 485  |
| QY                                                                          | 1809 | CAT-----CTGAACCTGCTAAACT-----CACATTCAGTATTCTTTTGTAGGATGTC     | 1855 |
| Db                                                                          | 484  | CATGATTTTAAAGTGAATGCAAGACTTTTAACTTTTAACTTTTCTAGGCATGTC        | 425  |
| QY                                                                          | 1856 | TGGGTTGTTGATTTATAGCAACTATGATATGAAGTACCAGTAACTTTTATATCTTT      | 1915 |
| Db                                                                          | 424  | AGGGTTTGTGGATATAGCAACTATGATATGAAGTACCAGTAACTTTTATATCTTT       | 365  |
| QY                                                                          | 1916 | GCAACGCAATGTTCTCTGCACTTATGCTTAGACTGCTTTTGTTCATAGTATACCGAGC    | 1975 |
| Db                                                                          | 364  | GCTGTGCTCTAGTACCTT-----TTCAAGTCTCTATGCCCCA                    | 329  |
| QY                                                                          | 1976 | TGAATTTATCTTCTCGAGGCGAGTGTGGATCTTTGATGTTCCCTTAAAAATTTTGTATG   | 2035 |
| Db                                                                          | 328  | CAATTTAATCTTCTTGGAGCCAGTGTGGGCTCTTGTATGTTCCCTTAAATTTTCTTGT    | 269  |
| QY                                                                          | 2036 | TGACAGATAAGAACTTTGATGCCCACTGAATTTGAAATGCTTTCTCTAGTCTTATATA    | 2095 |
| Db                                                                          | 268  | TACAGATAAGAACTTTGATGCTCTAAGTTTCAAAATGCTTCTCTAGTCTTATATA       | 209  |
| QY                                                                          | 2096 | CGGTTATGTTGATGCTTTCTTTGATTTGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA    | 2155 |
| Db                                                                          | 208  | CGGTTATGTTGATGCTTTCTTTGATTTGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA    | 150  |
| QY                                                                          | 2156 | TGAATTTACTGTTCTAGGTGAGGGAATATGATCCAGGAGTGTGAGTCCGAAGCCAGATGA  | 2215 |
| Db                                                                          | 149  | TAACTCTACTGTTTAGGTGAGGGAATATGATCCAGGAGTGTGAGTCCGAAGCCAGATGA   | 90   |
| QY                                                                          | 2216 | TTCTAAAAGCTATAGAAGCAGGAGTCCGAGCGCTGCTCAAGCT                   | 2259 |
| Db                                                                          | 89   | TTCTAGAAGCAGGAGTCCGAAGCTATAGTTATAGTAGCAGAAGGT                 | 46   |
| AL950061 298 bp DNA linear GSS 24-OCT-2002                                  |      |                                                               |      |
| Arabidopsis thaliana T-DNA flanking sequence GK-325E11-015977,              |      |                                                               |      |
| genomic survey sequence.                                                    |      |                                                               |      |
| ACCESSION AL950061                                                          |      |                                                               |      |
| VERSION AL950061.1 GI:24406683                                              |      |                                                               |      |
| KEYWORDS GSS.                                                               |      |                                                               |      |
| SOURCE Arabidopsis thaliana (thale cress)                                   |      |                                                               |      |
| ORGANISM Arabidopsis thaliana                                               |      |                                                               |      |
| Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;          |      |                                                               |      |
| Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;                |      |                                                               |      |
| rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.                 |      |                                                               |      |
| 1                                                                           |      |                                                               |      |
| REFERENCE Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H. |      |                                                               |      |
| AUTHORS and Weissnaar,B.                                                    |      |                                                               |      |
| TITLE A pipeline for automated high-throughput generation of FSTs           |      |                                                               |      |
| (flanking sequence tags) from Arabidopsis thaliana T-DNA                    |      |                                                               |      |
| transformed lines                                                           |      |                                                               |      |
| JOURNAL Unpublished                                                         |      |                                                               |      |
| REFERENCE 2                                                                 |      |                                                               |      |



TITLE Arabidopsis thaliana Gene Expression MicroArray  
JOURNAL Unpublished  
COMMENT Contact: David Smoller, Ph.D.  
Genome Systems, Inc., a wholly owned subsidiary of Incyte  
Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genomesystems.com.

FEATURES Location/Qualifiers  
source 1..439  
    /organism="Arabidopsis thaliana"  
    /mol\_type="mRNA"  
    /cultivar="Columbia Col-0"  
    /db\_xref="taxon:3702"  
    /clone="70167272"  
    /tissue\_type="rosette"  
    /dev\_stages="4 - 7 weeks"  
    /clone\_lib="A. thaliana, Columbia Col-0, rosette-1"  
    /note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."  
BASE COUNT 154 a 91 c 81 g 113 t  
ORIGIN

Query Match 5.6%; Score 227.4; DB 9; Length 439;  
Best Local Similarity 99.2%; Pred. No. 2.4e-23;  
Matches 239; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 3219 AAAGTTTGGACTATGATTCGATTCCTCAATATATTTATCTTTTGACATAGTGG 3278  
DB 439 AAAGTTTGGACTATGATTCGATTCCTCAATATATTTATCTTTTGACATAGTGG 380  
QY 3279 ATTCTGTCTTGAAGTCTTTCTAGGACAGCATTTAAGCTCCGGGACTAGATGGAGATG 3338  
DB 379 ATTCTGTCTTGAAGTCTTTCTAGGACAGCATTTAAGCTCCGGGACTAGATGGAGATG 320  
QY 3339 GTCAGTAAATTTCTTTGTTATGCCACACTTACATGGGGTTTTTCGGTCTTGTCGAGTTC 3398  
DB 319 GTCAGTAAATTTCTTTGTTATGCCACACTTACATGGGGTTTTTCGGTCTTGTCGAGTTC 261  
QY 3399 CCAATCAAGATCAAAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 3458  
DB 260 CCAATCAAGATCAAAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 201  
QY 3459 A 3459  
DB 200 A 200

Search completed: January 29, 2004, 02:49:17  
Job time : 7856 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 16:39:25 ; Search time 216 Seconds

Title: US-10-014-927-18  
Perfect score: 4044  
Sequence: 1 agacaaagatgcttacttct.....atgattatccttaactaac 4044  
8963.673 Million cell updates/sec

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length  | DB ID | Description         |
|------------|-------|-------------|---------|-------|---------------------|
| C 1        | 78.6  | 1.9         | 5285    | 4     | US-09-402-328-1     |
| C 2        | 75.8  | 1.9         | 7218    | 1     | US-08-232-463-14    |
| C 3        | 54.4  | 1.3         | 7218    | 1     | US-08-232-463-14    |
| C 4        | 48.8  | 1.2         | 193     | 4     | US-09-702-705-1719  |
| C 5        | 48.8  | 1.2         | 193     | 4     | US-09-736-457-1719  |
| C 6        | 48.2  | 1.2         | 284     | 4     | US-09-313-294A-5971 |
| C 7        | 47.8  | 1.2         | 19124   | 2     | US-08-487-826B-13   |
| C 8        | 45.8  | 1.1         | 2447    | 2     | US-09-014-959-14    |
| C 9        | 41.8  | 1.0         | 1117    | 3     | US-09-247-373B-33   |
| C 10       | 41.4  | 1.0         | 1586    | 4     | US-09-152-060-12    |
| C 11       | 41.4  | 1.0         | 2773    | 4     | US-09-996-243-178   |
| C 12       | 40.8  | 1.0         | 5511    | 3     | US-08-928-361B-2    |
| C 13       | 40.8  | 1.0         | 5511    | 4     | US-09-588-995A-2    |
| C 14       | 40.8  | 1.0         | 7334    | 3     | US-08-928-361B-1    |
| C 15       | 40.8  | 1.0         | 7334    | 4     | US-09-588-995A-1    |
| C 16       | 40.6  | 1.0         | 658     | 3     | US-08-998-416-595   |
| C 17       | 40.6  | 1.0         | 5163    | 3     | US-08-700-651-1     |
| C 18       | 40.6  | 1.0         | 5163    | 3     | US-08-928-361B-4    |
| C 19       | 40.6  | 1.0         | 5163    | 4     | US-09-588-995A-4    |
| C 20       | 40.6  | 1.0         | 5318    | 3     | US-08-700-651-2     |
| C 21       | 40.6  | 1.0         | 5318    | 3     | US-08-928-361B-3    |
| C 22       | 40.6  | 1.0         | 5318    | 4     | US-09-588-995A-3    |
| C 23       | 40.4  | 1.0         | 1798    | 4     | US-09-797-906-1     |
| C 24       | 39.6  | 1.0         | 665     | 3     | US-08-998-416-937   |
| C 25       | 39.6  | 1.0         | 1664976 | 4     | US-08-916-421B-1    |
| C 26       | 39.4  | 1.0         | 240     | 3     | US-08-628-417-6     |
| C 27       | 39.2  | 1.0         | 6755    | 3     | US-08-931-999-4     |

|      |      |     |         |   |                    |                    |
|------|------|-----|---------|---|--------------------|--------------------|
| 28   | 39   | 1.0 | 1664976 | 4 | US-08-916-421B-1   | Sequence 1, Appli  |
| C 29 | 38.6 | 1.0 | 662     | 3 | US-08-998-416-185  | Sequence 185, App  |
| C 30 | 38.6 | 1.0 | 701     | 3 | US-08-998-416-701  | Sequence 701, App  |
| C 31 | 38.6 | 1.0 | 711     | 3 | US-08-998-416-786  | Sequence 786, App  |
| C 32 | 38.6 | 1.0 | 724     | 3 | US-08-998-416-683  | Sequence 683, App  |
| C 33 | 38.6 | 1.0 | 722     | 3 | US-08-998-416-1036 | Sequence 1036, Ap  |
| C 34 | 38.6 | 1.0 | 767     | 3 | US-08-998-416-472  | Sequence 472, App  |
| C 35 | 38.6 | 1.0 | 782     | 3 | US-08-998-416-224  | Sequence 224, App  |
| C 36 | 38.6 | 1.0 | 827     | 3 | US-08-998-416-535  | Sequence 535, App  |
| C 37 | 38.6 | 1.0 | 828     | 3 | US-08-998-416-538  | Sequence 538, App  |
| C 38 | 38.6 | 1.0 | 834     | 3 | US-08-998-416-305  | Sequence 305, App  |
| C 39 | 38.2 | 0.9 | 4395    | 2 | US-08-849-480A-3   | Sequence 3, Appli  |
| C 40 | 38   | 0.9 | 494     | 2 | US-08-332-766A-22  | Sequence 22, Appli |
| C 41 | 38   | 0.9 | 1696    | 4 | US-09-835-811-1    | Sequence 1, Appli  |
| C 42 | 38   | 0.9 | 5852    | 1 | US-07-867-106-2    | Sequence 2, Appli  |
| C 43 | 38   | 0.9 | 6243    | 2 | US-09-056-075-1    | Sequence 1, Appli  |
| C 44 | 37.8 | 0.9 | 10640   | 4 | US-09-417-485D-5   | Sequence 5, Appli  |
| C 45 | 37.4 | 0.9 | 289     | 3 | US-09-007-005-17   | Sequence 17, Appli |

ALIGNMENTS

RESULT 1  
US-09-402-328-1/c  
; Sequence 1, Application US/09402328  
; Patent No. 6365728  
; GENERAL INFORMATION:  
; APPLICANT: Purdue Research Foundation,  
; Hodges, Thomas K.  
; Lysnik, Leszek A.  
; TITLE OF INVENTION: Regulatory Element For Expressing Genes  
; IN PLANTS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Barnes & Thornburg  
; STREET: 11 S. Meridian  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46204  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/402,328  
; FILING DATE: 05-No. 6365728-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Breen, John P.  
; REGISTRATION NUMBER: 38,833  
; REFERENCE/DOCKET NUMBER: 3220-29933  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 231-7745  
; TELEFAX: (317) 231-7433  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5285 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Arabidopsis thaliana  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-402-328-1

Query Match 1.9%; Score 78.6; DB 4; Length 5285;  
Best Local Similarity 51.1%; Pred. No. 5.8e-11;

STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463

FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/935,313  
 FILING DATE:  
 APPLICATION NUMBER: EP 91 114 300.6  
 FILING DATE: 26-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 CLONE: pTZgpt-F15  
 US-08-232-453-14

[illegible]

RESULT 3  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:

APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: BP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-F15  
US-08-232-463-14

Query Match 1.3%; Score 54.4; DB 1; Length 7218;  
Best Local Similarity 3.3%; Pred. No. 0.00024;  
Matches 13; Conservative 224; Mismatches 155; Indels 0; Gaps 0;  
QY 81 CTGCAAGCAACAGTGGAGATGATTGATTAACGGAATATCGGAACATCACTCAACAAC 140  
Db 1435 CRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1376  
QY 141 CAAAAATTGGACATCATATCGCAACAAATTCATAGGAAATACTGAAATTCACAAAC 200  
Db 1375 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1316  
QY 201 AGAAAAACCAACGACAGACAGACTACGAGTACGAGACACCGTGGACGGTGC 260  
Db 1315 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1256  
QY 261 ACGCGGAAAAATGCTTGGACGAGTACTAATCGCGAATTGAGATTGAGAGTGGT 320  
Db 1255 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1196  
QY 321 AGTAGAGCAACGAGAGATGTTCTCTCAAAAAATCCCAAGTGTTCGATCTAGT 380  
Db 1195 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1136  
QY 381 TCTCTTTTGCACAAACGACAGTGTTCAGGAACTAGGAAATGATGACCGCATGTT 440  
Db 1135 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1076  
QY 441 CGAATCCGATCGAAATGGTTCCAAATGTA 472

Db 1075 .RRRRRRRRATCGAAGCTCCCTCGACCTGCA 1044  
RESULT 4  
US-09-702-705-1719  
Sequence 1719, Application US/09702705  
Patent No. 6504010  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C14  
CURRENT APPLICATION NUMBER: US/09/702,705  
CURRENT FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 1833  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1719  
LENGTH: 193  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-702-705-1719  
Query Match 1.2%; Score 48.8; DB 4; Length 193;  
Best Local Similarity 77.6%; Pred. No. 0.0011;  
Matches 59; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 1172 TTGAAGATCCTCGTGTGATGACAGATGCAATTTATGACGTGATGTTATGATTTGATG 1231  
Db 1 TCGAGGACCCCGAGATGCGAGGATGCTATTATGGAAGAAATGTTATGATTATGCGC 60  
QY 1232 GGTGTCGACTTCGGGT 1247  
Db 61 AGTGTGCGTTCGTGT 76  
RESULT 5  
US-09-736-457-1719  
Sequence 1719, Application US/09736457  
Patent No. 6509448  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, Aijun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C15  
CURRENT APPLICATION NUMBER: US/09/736,457  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1719  
LENGTH: 193  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-736-457-1719  
Query Match 1.2%; Score 48.8; DB 4; Length 193;

|    | Best Local Similarity | 77.6%;                                                     | Pred. No. 0.0011; |     |            |
|----|-----------------------|------------------------------------------------------------|-------------------|-----|------------|
|    | Matches               | 59;                                                        | Conservative      | 0;  | Mismatches |
|    |                       |                                                            |                   | 17; | Indels     |
|    |                       |                                                            |                   |     | 0;         |
|    |                       |                                                            |                   |     | Gaps       |
|    |                       |                                                            |                   |     | 0;         |
| QY | 1172                  | TTGAGATCCTCGTGATGCAGACGATGCAATTTATGACGTCGATGGTTATGATTTCATG | 1231              |     |            |
| Db | 1                     | TCGAGACCCCGAGATGCAGAGGATGCTATTTTCGAGAGAAATGGTTATGATTATGCC  | 60                |     |            |
| QY | 1232                  | GGTGTGCACTTCGGGT                                           | 1247              |     |            |
| Db | 61                    | AGTGTGGGCTTCGTGT                                           | 76                |     |            |

## RESULT 6

```

US-09-313-294A-5971
/ Sequence 5971, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Lalgudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PE-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 5971
/ LENGTH: 284
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700350992H1
/ NAME/KEY: unsure
/ LOCATION: 70, 83, 238, 258-259, 283
/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5971

```

|    | Query Match           | 1.2%                                                        | Score 48.2;       | DB 4;     | Length 284; |
|----|-----------------------|-------------------------------------------------------------|-------------------|-----------|-------------|
|    | Best Local Similarity | 68.4%;                                                      | Pred. No. 0.0019; |           |             |
|    | Matches 65;           | Conservative 0;                                             | Mismatches 30;    | Indels 0; | Gaps 0;     |
| QY | 1153                  | GTTTGTTATTGGTGGCGAGTTTGAAGATCCCTCGTGAATGCAGATGCATTTATGGAGCT | 1212              |           |             |
| Db | 24                    | GCTTATGCTTTTGTGAGTTTGAAGATCCCTCGTGAATGCAGAGGNAATTGCTGAACGN  | 83                |           |             |
| QY | 1213                  | GATGTTATGATTTTGAATGGGTGTCGACTTCGGGT                         | 1247              |           |             |
| Db | 84                    | GATGATACAACTTTGATGCACACCGCTCAAGAGT                          | 118               |           |             |

## RESULT 7

US-08-487-826B-13/C  
 ; Sequence 13, Application US/08487826B  
 ; Patent No. 5993827  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sim, Kim L.  
 ; APPLICANT: Chitnis, Chetan  
 ; APPLICANT: Miller, Louis H.  
 ; APPLICANT: Peterson, David S.  
 ; APPLICANT: Su, Xin-zhaun  
 ; APPLICANT: Wellemis, Thomas E.  
 ; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 ; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Knobbe Martens Olson & Bear  
 ; STREET: 520 Newport Center Drive 16th Floor  
 ; CITY: Newport Beach  
 ; STATE: California  
 ; COUNTRY: US  
 ; ZIP: 92660  
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PS-08-487-826B-13

|                       |              |                  |                 |                   |
|-----------------------|--------------|------------------|-----------------|-------------------|
| Query Match           | 1.2%;        | Score 47.8;      | DB 2;           | Length 19124;     |
| Best Local Similarity | 44.0%;       | Pred. NO. 0.026; |                 |                   |
| Matches 202;          | Conservative | 0;               | Mismatches 257; | Indels 0; Gaps 0; |

|    |       |                                                                 |       |
|----|-------|-----------------------------------------------------------------|-------|
| Qy | 2449  | CTGTGCTCAAGGTATGAGTGTAGATTTGTATCATTTATATATATGTAGTTACCCCTTCA     | 2508  |
| Db | 15862 | TTTTTTTATTTAAATAAATTTTTTTTATTTATTTATTTTTTTTATTTAAATAAATTTTTT    | 15803 |
| Qy | 2509  | TGGATCACATGTGTTCTGTGCATAGTGAACCTCTACTAGCTTTATTACTTACAACTAAGCAC  | 2568  |
| Db | 15802 | TTATTTTATGTATATATTTTTTTTTTTTAAACAATTTTTTAAATTTTTTTTATTTTATGATAT | 15743 |
| Qy | 2569  | CTTTTGTTCGCTTCGGTACACAGTTCGAATTTGTTTGAGTCTTTTTTCCCTCATAGTGGACT  | 2628  |
| Db | 15742 | ATATTTTAAATTAATAATATTTTTTTCTTTTTTTTTTTTATGATATATATTTTTTT        | 15683 |
| Qy | 2629  | AGTCATATCTCACTGATTTTCTCTCTTGTTGTAGTGTTCCTATGTCATGCACAACTCCAA    | 2688  |
| Db | 15682 | TTTTTTAAAGTTTTTTTTTTTTCTCTCTTGTTTTTATTTTTTTTAAATCAATTTTTTTTT    | 15623 |
| Qy | 2689  | TATGGTAAAGTGTACCTCCTTGTGTTGGGATTTACCAGAGTTCCTTTTCATTTCTTTACAG   | 2748  |
| Db | 15622 | TATATAAAATTTTTTTTAAATTTTTTTTGATAATCTTTTTTCATTTTTTATCTATCAAA     | 15563 |
| Qy | 2749  | TGAATCGTGTCTGTTTTTATGTTTTGAGTTCCTTGACAGAGATGCTCCCATCATATTAGT    | 2808  |
| Db | 15562 | ATTATATTTTATTAATAATTTTTTATTTTTTAAAAAATTTTTTCTCTTTTTTTTTTTTT     | 15503 |
| Qy | 2809  | CGTTTTCTCTTCTTTGTGCGTCTCTCTCTGGAGTTCCTCTGATGAAGGCTTTACT         | 2868  |
| Db | 15502 | TTTTTTTTTATTTTAAATAAATTTTTTTTTTTATTTTCATTTTTTCTTTTTTCATTTTAATA  | 15443 |
| Qy | 2869  | TCCTAACTTTTTTCCAGCGACGGTGAATTTATTACGTAT                         | 2907  |
| Db | 15442 | AAATTCCTTTTTTATATCTCTTTTTTAAATAATAAATACATAT                     | 15404 |

## RESULT 8

US-09-014-969-14/C  
; Sequence 14, Application US/09014969  
; Patent No. 5965397  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaVallee, Edward R.  
; APPLICANT: Racie, Lisa A.  
; APPLICANT: Merberg, David

```

RESULT 9
US-09-247-373B-33/c
; Sequence 33, Application US/09247373B
; Patent No. 6166954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1108-A
; CURRENT APPLICATION NUMBER: US/09/247,373B
; CURRENT FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/524,747
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97

```

```

RESULT 10
US-09-152-060-12/c
; Sequence 12, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003Pl.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens

```





[illegible]

QY 2250 GGTCCAAAGCTAGCTATAGTAGCAAGCAGGAGGCTCTTTTTTTTTTTTTTTTTTTTTCAT 2309  
Db 978 TGTGTTGTTGTTGTCGTAGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTTGTT 919  
QY 2310 AAACCTAAGCATATAGGAGTTTATTTGTTAACTTTATGAATTAAT 2361  
Db 918 CACTCCACACATTTAGTGTATGAATGTATACCTTACTGCCCATATCAAT 867

RESULT 13  
US-09-588-995A-2/c  
; Sequence 1, Application US/09588995A  
; Patent No. 6514697  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: BARNES, DEBRA A.  
; APPLICANT: NELSON, RICHARD C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND  
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-5  
; CURRENT APPLICATION NUMBER: US/09/588,995A  
; PRIOR FILING DATE: 2000-06-06  
; PRIOR FILING DATE: 1997-03-27  
; PRIOR FILING DATE: 1997-03-27  
; PRIOR FILING DATE: 1997-09-12  
; PRIOR FILING DATE: 1996-08-14  
; PRIOR FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 115  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 5511  
; TYPE: DNA  
; ORGANISM: Cryptosporidium parvum  
US-09-588-995A-2

Query Match 1.0%; Score 40.8; DB 4; Length 5511;  
Best Local Similarity 52.3%; Pred. No. 1;  
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 2190 AGGAGTGTGAGTCGAAGCCAGATCTCTAAAGCTATAGAGCAGGAGTGGAGCCGT 2249  
Db 1038 AGTAGTTGTTGTCGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTTGTTGTT 979

QY 2250 GGTCCAAAGCTAGCTATAGTAGCAAGCAGGAGGCTCTTTTTTTTTTTTTTTTTTTTTCAT 2309  
Db 978 TGTGTTGTTGTTGTCGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTTGTTGTT 919

QY 2310 AAACCTAAGCATATAGGAGTTTATTTGTTAACTTTATGAATTAAT 2361  
Db 918 CACTCCACACATTTAGTGTATGAATGTATACCTTACTGCCCATATCAAT 867

RESULT 14  
US-08-928-361B-1/c  
; Sequence 1, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto

STATE: CA  
COUNTRY: USA  
ZIP: 94306-1840  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,361B  
FILING DATE: 12-SEP-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,062  
FILING DATE: 13-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Verny, Hana  
REGISTRATION NUMBER: 30,518  
REFERENCE/DOCKET NUMBER: 480.76-1 (HV)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-1677  
TELEFAX: 650-324-1678  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7334 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-928-361B-1

Query Match 1.0%; Score 40.8; DB 3; Length 7334;  
Best Local Similarity 52.3%; Pred. No. 1.2;  
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 2190 AGGAGTGTGAGTCGAAGCCAGATCTCTAAAGCTATAGAGCAGGAGTGGAGCCGT 2249  
Db 2706 AGTAGTTGTTGTCGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTTGTTGTT 2647

QY 2250 GGTCCAAAGCTAGCTATAGTAGCAAGCAGGAGGCTCTTTTTTTTTTTTTTTTTTTTTCAT 2309  
Db 2646 TGTGTTGTTGTTGTCGTAGTAGTAGTAGTTGTTGTTGTTGTTGTTGTTGTTGTT 2587

QY 2310 AAACCTAAGCATATAGGAGTTTATTTGTTAACTTTATGAATTAAT 2361  
Db 2586 CACTCCACACATTTAGTGTATGAATGTATACCTTACTGCCCATATCAAT 2535

RESULT 15  
US-09-588-995A-1/c  
; Sequence 1, Application US/09588995A  
; Patent No. 6514697  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: BARNES, DEBRA A.  
; APPLICANT: NELSON, RICHARD C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND  
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-5  
; CURRENT APPLICATION NUMBER: US/09/588,995A  
; PRIOR FILING DATE: 2000-06-06  
; PRIOR FILING DATE: 1997-03-27  
; PRIOR FILING DATE: 1997-03-27  
; PRIOR FILING DATE: 1997-09-12  
; PRIOR FILING DATE: 1996-08-14  
; PRIOR FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 115  
; SOFTWARE: Patent in Ver. 2.1

```
; SEQ ID NO 1
; LENGTH: 7334
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-1

Query Match      1.0%; Score 40.8; DB 4; Length 7334;
Best Local Similarity 52.3%; Pred. No. 1.2;
Matches 90; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY      2190 AGGAGTGTGAGTCGAGCCCGAGATGATTCCTAAAGCTATAGAACGAGGAGTCGGAGCCCT 2249
         |||||
Db       2706 AGTAGTTGTTGTTGTCCTAGTAGTAGTGGTAGTAGTTGTAGTAGTTGTGTTGTCGT 2647
         |||||

QY      2250 GGTCCAAGCTGTAGCTATAGTCAGAGCAGGAGGCTTTTTTTTTTTTTTTTTCAT 2309
         |||||
Db       2646 TGTGTTGTTGTTGTCCTAGTAGTAGTAGTTGTGTTGTTGTTGTTGTTGTTGTT 2587
         |||||

QY      2310 AAACCTAAGACATATAAGGATTTTATTGTAACCTTATTATGAAATTAAT 2361
         |||||
Db       2586 CACTCCAACACATTTAGTGTATGGAATGTATACCTTACTGCCCATATCAAT 2535
         |||||

Search completed: January 29, 2004, 00:38:13
Job time : 219 secs
```



Creation date: 04-12-2004  
Indexing Officer: SMOHAMMED - SUAD MOHAMMED  
Team: 1600PrintWorkingFolder  
Dossier: 10133973

Legal Date: 04-09-2004

| No. | Docode | Number of pages |
|-----|--------|-----------------|
| 1   | CTRS   | 7               |

Total number of pages: 7

Remarks:

Order of re-scan issued on .....